Use of Prediction Models to Investigate Respiratory Support Therapy of Infants with Acute Viral Bronchiolitis:
Retrospective Observational Study using Machine-Learning Techniques at a large Tertiary Center

Christoph Camphausen MD, PhD

Conflict of Interest Statement

The author declares that there are no conflicts of interest.

Statement of Candidate

This thesis is presented as a fulfillment to the requirements for the degree Masters of Philosophy.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree to any other university or institution other than Macquarie University.

Approval for the research presented in this thesis was obtained from the Macquarie University Human Research Ethics Committee (Reference Number: 5201700411). De-identified clinical records used in this research were provided by the Children's Hospital of Los Angeles, USA.

I also certify that this thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself has been appropriately acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Dr Christoph Camphausen (MQ ID:)

18 December 2019

Ethics approval for research

Ethics application reference number 5201700411 was reviewed by the Macquarie University Human Research Ethics Committee (HREC (Medical Sciences)).

Documents reviewed	Version no.	Date
Macquarie University Human Research Ethics Committees Prior Review Form (PREF)	N/A	Received 10 Apr 2017

Documents Noted	Version no.	Date
Notice of IRB Exemption – with stipulations (Children's Hospital Los Angeles Institutional Review Board)	N/A	13 Apr 2017
IStar Application (Children's Hospital Los Angeles)	1.3	12 Apr 2017
Research Protocol	2.3	11 Apr 2017
Data Collection Sheet	2.3	11 Apr 2017
Data Use Agreement- Protected Health Information	N/A	Signed 24 Feb 2017

This research was found to meet the requirements set out in the National Statement on Ethical Conduct in Human Research (2007 - Updated March 2014) (the National Statement) and approval was provided on the 2nd of May, 2017.

USC Children's Hospital LOS ANGELES. LAC+USC HEALTHCARE NETWORK   **iStar**

May 03, 2017, 07:23am

To: Christoph Camphausen
ACCM Research Oversight Committee - CHLA
Jeni Kwok, JD
ACCM Research Oversight Committee - CHLA
Christopher Newth
ACCM Research Oversight Committee - CHLA

From: Dongqing Li, M.B., M.B.S.
Chair Designee, Children's Hospital Los Angeles Institutional Review Board

Re: CHLA-17-00117   Christoph Camphausen
Use of Prediction Models to Optimize Respiratory Support Therapy of Infants with Acute Viral Bronchiolitis: Retrospective Observational Study using Machine-Learning Techniques at a large Tertiary Center

NOTICE OF APPROVAL OF AMENDMENT - WITH STIPULATION*
(Reference: Bronchiolitis Study - Macquarie Univ Approval - CHLA-17-00117-AM001)
Review Date: 5/2/2017

***Stipulation: Only data existing between January 1, 2008 through March 26, 2017 may be used for the purpose of this study.**

Document(s) Reviewed:
- iStar Application with Amendment Application AM001 (Dated: 5/2/2017)

The amendment was reviewed and approved by a member of the CHLA IRB. The study is still exempt per 45 CFR 46.101[b][#4].

Acknowledgement

Firstly, I would like to thank my excellent supervisors: Professor Enrico Coiera, for his expertise, wisdom, and direction; and his reliable assistance until the final submission of my thesis. I would also like to thank Associate Professor Blanca Galled Luxan, for her expedient efforts, time, and guidance through the complexities of clinical analytics, machine learning, and research design.

I would also like to express my sincere appreciation to Professor Christopher Newth and Dr. Matthew Keefer at the Children's Hospital of Los Angeles for the opportunity to work, for welcoming me as an external researcher, and sharing their workspace with me. Without their generous support and effort, I would have not gained research access to the hospital's clinical database and I am ever grateful.

Finally, I would like to sincerely thank my family and friends for their love and patience as I entered the world of machine learning and causal inference. In particular, I want to thank my wife, Chana, for her loving and motivational support. I could not have finished it without her.

Abstract

Acute viral bronchiolitis (AVB) is the most common lower respiratory tract infection during the first year of life and the most frequent reason for hospitalization during infancy, generating extensive cost for healthcare systems.

The overall aim of this thesis was to understand and optimise current respiratory therapies for AVB patients who presented to a large tertiary hospital in the USA. The study design was retrospective and observational. It used machine-learning techniques and causal inference algorithms to inform clinical decision making. Specifically, it compared the effectiveness of high-flow nasal cannula (HFNC) with that of standard treatment. The primary outcome was length of hospital stay.

The dataset contained all AVB patients under the age of one year who presented to the Children's Hospital of Los Angeles between 01/2008 and 03/2017. In total, 891 patients were admitted and treated with either standard nasal cannula therapy or HFNC. The dataset was reduced to 599 cases after excluding significant co-morbidities and outliers to ensure the study was truly focussing on AVB patients.

The analysis was performed in four steps: descriptive statistics, feature selection, data visualisation, propensity score matching, and predictive analytics. Propensity score matching was used to match patients in the standard group with those in the high flow group. Subsequent regression analysis estimated the average treatment effect of HFNC on the primary outcome.

Due to decision bias, propensity score matching could not demonstrate a treatment effect of high flow therapy on hospital length of stay. This finding was in accordance with the latest literature.

The list of the examined confounding variables included patient demographics, common co-morbidities, viral cause, vital parameters, and clinical descriptors of the respiratory state of the patient. In total, the combined influence of 22 covariates on the treatment choice and outcome was investigated. A newly created data-driven respiratory severity score incorporated those 22 covariates and converted them into individual scores. The sum of the individual scores generated a respiratory severity score.

Respiratory severity scores, obtained at different times of the hospital stay, and other covariates (risk factors) were used to fit machine learning models that predicted hospital length of stay, prolonged length of stay (>5 days), the need for high flow therapy, and failure of standard therapy or high flow therapy. The results were highly significant. In addition to face and content validity, construct and prediction validity were successfully evaluated by applying statistical and machine learning tools.

The respiratory severity score demonstrated promising characteristics when used in a fully computerised healthcare setting. As soon as full validation is achieved, it has the potential to become a useful instrument for clinical decision making, randomized controlled trials and comparative effectiveness research.

Summary of Results

Single centre, retrospective, observational EHR study

PREDICTION OF

HOSPITAL LENGTH OF STAY	
Highest-ranking individual parameters of respiratory distress	- capillary pO_2 - signs of retractions and respiratory effort - breath sounds - heart rate - body temperature - respiratory rate - pulse-oximetry
Severity score calculated from the second six-hour period following the first recorded event	$p<0.01$
Due to decision bias, propensity score matching could not demonstrate a treatment effect of high flow therapy	

PROLONGED LENGTH OF STAY (>5 days)	
Severity score measured at the time of the treatment decision to apply standard care or high flow therapy	$p<0.001$. OR 1.215
Corrected age at time of admission	OR 0.922
Z score of the body weight	OR 0.875 (97.5% CI; 0.816 to 0.936)
Acute viral bronchiolitis as the only diagnosis (simple case)	OR 0.403 (97.5% CI; 0.265 to 0.606)
Metapneumo virus	OR 2.322 (97.5% CI; 1.065 to 4.995)

HIGH FLOW THERAPY	
Severity score	ROC AUC 0.84, Naïve Bayes model
Severity score in conjunction with six other covariates	ROC AU 0.83, sensitivity 0.94, specificity 0.33. OR 2.8 (97.5% CI, 2.23 to 3.52) Generalized linear model (GLM)
Age younger than 100 days	OR 2.97 (97.5% CI; 1.737 to 5.19)
Z score of the body weight of less than -2	OR of 2.57 (97.5% CI; 1.175 to 5.47)
RSV, RhinoEntero virus, Metapneumo virus	Increased odds ratios

FAILURE OF STANDARD THERAPY OR HIGH FLOW THERAPY	
Severity score of first and second third of standard treatment	Negative correlation
Severity score of second third of high flow treatment	Negative correlation
High severity scores of 3rd third during high flow and standard treatment	Positive correlation
Increasing severity scores during high flow or standard treatment (=positive slope of regression line)	Positive correlation

OR = odds ratio

Table of Contents

List of Figures

List of Code Snippets

List of Tables

Definitions and Abbreviations

LOS   Length of Stay

HFNC   (Heated) High-flow oxygen therapy through a nasal cannula

HR   heart rate (beats per minute = bpm)

RR   respiratory rate (breaths per minute = bpm)

ED   emergency department

tLOS   total length of stay

CCL   Cerner Command Language

SQL   Structured Query Language

ICD   International Classification of Diseases

CSV   Comma separated Values

NLP   Natural language processing

HIPAA   Health Insurance Portability and Accountability Act

PHI   Protected Health Information

MRN   medical record number

DOB   date of birth

AVB   Acute Viral Bronchiolitis

LOS   Length of Stay

CCL   Cerner Command Language

SQL   Structured Query Language

Glossary

Data Definition Language (DDL) statements are used to define the database structure or schema.
```
e.g. CREATE, ALTER, DROP, TRUNCATE, COMMENT, RENAME
```

Data Manipulation Language (DML) statements are used for managing data within schema objects.
```
e.g. SELECT, INSERT, UPDATE, DELETE, MERGE, CALL, EXPLAIN PLAN, LOCK TABLE
```

HIPAA   Health Insurance Portability and Accountability Act of 1996) is United States legislation that provides data privacy and security provisions for safeguarding medical information, i.e. patient health information (PHI). **PHI** includes:
- a patient's name, address, birth date and Social Security number;
- an individual's physical or mental health condition;
- any care provided to an individual; or
- information concerning the payment for the care provided to the individual that identifies the patient, or information for which there is a reasonable basis to believe could be used to identify the patient.

Pulse oximetry is a non-invasive method for monitoring a person's oxygen saturation (SpO_2).
Normal values are between 95% and 100%.

Fraction of Inspired Oxygen (FiO_2) represents the percentage of oxygen participating in gas-exchange. Natural air includes 21% oxygen, which is equivalent to FiO2 of 0.21.

Supplemental oxygen means an FiO_2 greater than the 21% oxygen in room (ambient) air. When supplemental oxygen is applied to a patient, the patient's inhaled FiO_2 is increased above 21%, i.e. $FiO_2 > 0.21$; the highest FiO_2 possible is 1.0, which represents 100% oxygen.

Litre per Minute of Oxygen versus FiO2. Standard oxygen sources can deliver from ½ litre per minute (L/min) of oxygen (O_2) to 5 L/min). Every litre per minute of oxygen increases the percentage of O_2 the patient breathes by 3–4%. Room air has a FiO_2 of 0.21. For example, if a patient is on 4 L/min O_2 flow, then the FiO_2 is between 0.33 and 0.37. Flow rates of **1-4 litres** per minute are used with nasal cannulas, equating to a concentration of approximately 24-40% oxygen, i.e. FiO_2 0.24-0.4.

1   Introduction

Electronic health records have become nearly universal in documenting every aspect of clinical care. Large amounts of data are now readily available and accessible. This provides new opportunities for comparative effectiveness research which is defined by the U.S. Department of Health and Human Services (DHHS) as: "*conduct and synthesis of research comparing the benefits and harms of different interventions and strategies to prevent, diagnose, treat and monitor health conditions in 'real world' settings.*"[1,2]

This retrospective, observational study used advanced data processing and machine learning to investigate aspects of respiratory support therapy in infants with acute viral bronchiolitis. The raw dataset, obtained from the Children's Hospital Los Angeles, comprised over ten million clinical events. The lack of randomization in observational studies necessitated the collection of numerous variables that were maybe associated with treatment selection and patient outcome. The aim was to develop and validate prediction models that might assist with clinical decision making.

The specific research question was, how high flow oxygen application via nasal cannula compared with standard low flow oxygen application in the management of acute viral bronchiolitis in infants. What are the main determinants of treatment failure and outcome? Are the results of this study strong enough to replicate results of randomised controlled trials?

1.1   Background

1.1.1   Acute Viral Bronchiolitis

Acute viral bronchiolitis is the most common lower respiratory tract infection during the first year of life and one of the main reasons for hospitalization during infancy. In the USA, acute viral bronchiolitis accounts for up to 18% of hospitalizations annually, costing approximately $1.73 billion.[3-6]

Acute viral bronchiolitis is a self-limited viral disease with lower airway inflammation. Symptoms can vary greatly, ranging from mild congestion and cough to severe respiratory distress, apnoea, or hypoxemia. Approximately 5-9% of bronchiolitis patients require active respiratory support, usually applied as non-invasive ventilation, and approximately 2% require intubation and mechanical ventilation. The risk of mortality is low (<0.05%).[3,7,8]

1.1.2   Non-Invasive Ventilation

The use of non-invasive ventilation has mostly replaced conventional mechanical ventilation for the treatment of acute viral bronchiolitis.[9,10] Nasal continuous positive airway pressure (nCPAP) and low flow nasal oxygen have been the standard respiratory therapy for many years.[11] Approximately ten years ago, the application of heated and humidified oxygen via high-flow nasal cannula (HFNC) was introduced.[12-14] The latest research reported no effect on duration of oxygen

therapy or length of hospital stay. However, high-flow oxygen therapy had significantly lower failure rates than standard oxygen therapy.[15-18]

Based on a large dataset, several current research questions in relation to high-flow nasal cannula were investigated:[19,20]

1) Development of a severity score and predictive models for high-flow nasal cannula
2) Evaluation of high-flow nasal cannula in comparison to low flow nasal cannula
3) Identification of clinical parameters that predict failure of high-flow oxygen therapy
4) Estimation of the treatment effect of high flow nasal cannula on length of stay
5) Evaluate predictors of length of stay

1.1.3   Electronic Health Records

In the USA, the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 led to an almost universal adoption rate of electronic health records.[21,22] Secondary usage of EHR data has increasingly produced research output which in return has started the transition process of EHR systems from documentation and billing to a dynamic research and learning tool.[23-25]

1.1.4   Machine Learning and Observational Studies

Machine learning methods are being increasingly used in healthcare.[26-28] The introduction of powerful computers and sophisticated statistical software has enabled clinical scientists to conduct studies on observational, "real world" data. Furthermore, the results of observational studies have the potential to complement findings of randomized controlled trials. Observational studies sometimes represent the only research option, especially when it is unethical or infeasible to randomly assign treatment alternatives. [29-31]

This study utilized many machine learning tools. Propensity score matching helped to balance standard group and high flow group based on observed covariates. The aim was to create two cohorts of patients that were as similar as possible except for their exposure to treatment, i.e. high flow nasal cannula. Generalized linear models were used to apply linear and logistic regression to compare groups of covariates in relation to outcome and intervention. Multiple machine learning models were fitted and validated. Finally, performance parameters (e.g. ROC, sensitivity, specificity, odds ratio) were used to decide on the best predictors and best fitted model.

2 Literature Review

To assist with the initial design of the study, the author performed a systematic review of the current literature on acute viral bronchiolitis to identify knowledge gaps and current research areas. As a result, three topics were identified:

1) Development and estimation of clinical severity scores
2) comparing two different methods of non-invasive ventilation
3) individualized recommendations for NIV[14]

2.1 Search Strategy and Selection Process

The search comprised journal articles between 2010 and 2018 which were located on PubMed at the US National Library of Medicine and EMBASE (OVID interface). The search term for PubMed was "Bronchiolitis"[MH] AND "last 7 year"[DP] AND English[lang] AND "infant"[MH]. In addition, the terms "bronchiolitis" and "child" were used to find registered trials via ClinicalTrials.gov[32] and Australian New Zealand Clinical Trials Registry (ANZCTR).[33] The Cochrane Database of Systematic Reviews (CDSR) was used to find reviews by using terms related to acute viral bronchiolitis and child. The author also checked reference lists of relevant articles and searched conference abstracts and searched for literature dealing with data analysis and statistical learning in relation to acute viral bronchiolitis.

The inclusion and exclusion criteria were as follows:

Inclusion Criteria

- Child less than two years of age
- First episode of acute viral bronchiolitis
- English language
- Articles published in peer-reviewed journals from January 2010 to March 2018

Exclusion Criteria

- Severe congenital or acquired heart disease
- Chronic lung disease
- Disease affecting normal functioning of the lung and heart (e.g. cerebral palsy, syndromes e.g. Pierre Robin)
- Extreme Prematurity
- Studies with sample size less than 30 objects
- Studies without severity assessment and use of non-validated severity scores

Two independent reviewers manually selected the articles for final review. The focus was on studies producing the best possible evidence, together with systematic reviews and Cochrane reviews which provided quick access to current scientific knowledge.

The data extraction for articles that described clinical severity scores included the individual score items. These were evaluated and previously described quality criteria were used to create a quality rating.[34] With regards to non-invasive ventilation, relevant properties of observational and RCT studies were extracted as follows: general topic, author, country, year of publication, type of study, centre, setting, number of subjects, type of intervention, comparator, age statistics, time periods, outcomes, result.

2.2 Results

The total of the initial search was 786 articles which were manually reviewed and reduced to 172 articles. Out of these, 27 articles contained systematic reviews and Cochrane reviews. There were 36 articles providing practice guidelines, editorials, and commentaries. The remaining 109 articles were categorized into three main research areas, clinical severity score (23), inhalation therapy with hypertonic saline (45), and non-invasive respiratory support (41). The numerical summary of the 109 articles is shown in Table 1.

Table 1: Literature Review - Overview of Research Areas

Research Area	RCT	OBS	Other	Total	Knowledge Gap/Finding
Clinical Severity Score	2	11	10	23	Lack of fully validated CSS
Inhalation with Hypertonic Saline*	26	3	16	45	The evidence suggests no effect of HS
Non-invasive ventilation	9	8	24	41	Optimum timing and indication not established yet
				109	

RCT – randomized controlled trial, OBS – observational study, *HS usually 3% solution

Clinical severity scores were investigated in two randomized controlled trials (RCT) and eleven observational studies. Many of the severity scores were specifically designed for acute viral bronchiolitis. Inhalation therapy with hypertonic saline has been extensively studied since its introduction in 2003. The results of the 26 randomized controlled trials were summarized in multiple systematic reviews and Cochrane reviews. Most reviewers now agree that the effect of hypertonic saline inhalation is minimal and does not warrant further studies. Although inhalation of hypertonic saline was utilized at Children's Hospital Los Angeles, the retrospective design of this study did not allow for a comprehensive data analysis. Respiratory support in acute viral bronchiolitis, mainly non-invasive ventilation, was discussed in eight observational studies and nine randomized controlled trials. A further 24 articles (e.g. editorials, commentaries) delivered background information about current trends in non-invasive ventilation.

2.2.1 Clinical Severity Score

To better understand and describe disease severity, clinicians and researchers have recognized the need for a robust clinical severity score. Many different scores have been introduced over the last 30 years; most of them with little or no validation.[35]

Table 2: Overview of The Most Commonly Used Clinical Severity Scores For Bronchiolitis And Dyspnoea

Assessment Tool (chronological order)	Duration of symptoms [days]	General condition	Appearance	Feeding	Dyspnoea WOB↑[1]	Apnoea	Breath Sounds[2]	Wheeze	Retractions	Oxygen Need[3]	SaO2	Cyanosis	RR	HR	Urine Output	Cap Refill time	Dehydration (WHO)	Age	Bekhof Quality Criteria (max. 15)
RDAI and RACS, Lowell 1987							✔	✔				✔							3
Wang, 1992		✔					✔	✔				✔							3
Liu 2004				✔			✔	✔				✔							4
Walsh 2006 predict hospital admission								✔					✔			✔	✔		N/A
Gajdos, 2009							✔	✔				✔							4
Marlais 2011 predict hospital admission	✔									✔		✔	✔					✔	7
CHWRS - Weisgerber, 2011		✔	✔	✔			✔	✔	✔	✔		✔	✔						4
WARM Respiratory Score Children's Hospital Cincinnati						✔ (air exchange)		✔	✔			✔							N/A
m-WCAS Duarte-Dorado, 2013		✔ (cerebral function)				✔ (inspirat.)		✔ (expirat.)	✔	✔									5
Tal, 1983							✔	✔			✔	✔							4
modified Tal, 2013							✔	✔		✔	✔								5
ABSS - Ramos Fernández, 2015				✔ (I/E Ratio)			✔	✔	✔			✔[6]	✔[7]						5
Faber, 2015 PulmoTrack								✔											N/A
LIBBS-PRO Ver. 9.2 van Miert, 2015		✔	✔	✔	✔	✔				✔		✔[5]	✔[5]	✔	✔				11
BROSJOD Score 2017							✔	✔	✔	✔	✔	✔	✔						7

[1]Occasional breaks with feeds, Frequent breaks with feeds, Unable to feed; Complete sentences, Phrases, Single words; Minimal/Some/Significant ↑ WOB

[2]Rales/crackles, insp. and exp. Wheeze, Rhonchi/Coarse, Prolonged Exp., poor air entry

[3]Room air, fiO2, cannula, simple mask, oxygen L/min

[4]Reference values according to age range

[5]Two age ranges (infants <3 months and infants ≥3 months)

[6]Resp. Rate: age ranges <2 mon, 2-6 mon, 6-12 mon

[7]Heart Rate: age ranges 7 days to 2 mon, 6-12 mon

Table 2 depicts an overview of fifteen clinical severity scores found in this literature review, listing eighteen different properties. Approximately 61% of the properties can be measured accurately. The remaining 39% either rely on clinical skills (e.g. auscultation) or are not well defined (e.g. general condition, appearance). Bekhof et al. developed quality criteria for paediatric dyspnoea scores and outlined in a systematic review that only a few clinical severity score had reached a high degree of validation.[34,36-38] By applying the Bekhof criteria only three severity scores were found to have a rating higher than six: Marlais, BROSJOD and LIBSS-PRO.[39-41]

The Liverpool Infant Bronchiolitis Severity Score - Proxy Reported Outcome (LIBSS-PRO), was developed as part of a PhD thesis. Psychometric methods were used to develop the scoring instrument, followed by extensive testing of validity and reliability. The Bekhof quality rating appears to be as high as eleven. The LIBSS-PROs has not been fully validated for responsiveness to change and cross-cultural differences yet.[41]

The Bronchiolitis Score of Sant Joan de Déu (BROSJOD) was evaluated in 2017.[40] Many items of the BROSJOD require specialized auscultation skills and advanced knowledge of the score which reduces it usability. The authors claim a quality rating of fourteen by using the Bekhof criteria. However, it appears that many criteria were not met. Therefore, the author of this review applied a Bekhof rating of seven (see Appendix D for more details).

The Bronchiolitis Risk of Admission Scoring System was developed by logistic regression analysis of case notes. The five best predictors of admission (age, respiratory rate, heart rate, oxygen saturations and duration of symptoms) were incorporated into the score which achieved a Bekhof quality rating of seven (validity 2, utility 5). Its simplicity makes automated analysis straight forward when using an EHR system. However, it can only be used to predict admission to hospital.[39]

2.2.2   Non-invasive Ventilation (NIV)

Several types of non-invasive ventilation are available: high flow nasal cannula, nasal continuous positive airway pressure (nCPAP), low flow nasal oxygen, headbox, and helmet.

This review identified eight relevant observational studies between 2010 and 2018 that dealt with aspects of non-invasive ventilation (Table 3, page 22).

Table 3: Overview Observational Studies Investigating Non-Invasive Ventilation

Topic	Author	Country	Published Year-Mon	Type	Centre	Setting	N	Max Age/ Median [month]	Time Period 1		Time Period 2		Outcomes	Results	Comments
									Start	End	Start	End			
HFNC Ward	Riese[17]	USA	2017-05	R	S	Ward	576	12/4.5	2010	2012	2012	2014	LOS Secondary: PICU LOS, PICU transfer rate, IV, 30-day re-admission rate	Use of HFNC increased from 24% to 35%. No effect after Introducing HFNC to peripheral wards.	APR-DRG severity level used as CSS. LOS was measured in days (not hours)
	Riese[42]	USA	2015-12	R	S	Ward, PICU	290	15/8.5	2010	2012	2012	2014	LOS Secondary: hosp. charges, IV, 30-day re-admission rate	Reduced LOS and hospital charges for bronchiolitis patients initially admitted to the PICU after Introducing HFNC to peripheral wards.	APR-DRG severity level used as CSS. LOS was measured in days (not hours)
HFNC Failure	Abboud[43]	USA	2012-11	R	S	PICU	113	12	2006	2010			HFNC failure	Non-responders: higher pCO2 and less tachypnoeic than responders; RR unchanged; PRISM-III high	
HFNC Success	McKiernan[44]	USA	2010-04	R	S	PICU	115	24/3	2005	2006	2006	2007	Intubation rate, PICU LOS	After introduction of HFNC 9% intubation vs. 23% prior to introduction of HFNC. PICU LOS decreased from 6 to 4 days.	
CPAP	Evans[45]	UK	2011-09	R	S	ED	163	12	2009	2010			Need for nCPAP	Strongest predictors were FiO2, RR, HR, age, gestational age, SaO2, GCS	
CPAP – CMV	Essouri[46]	France	2013-10	R	S	PICU	525	4	1996	2000	2006	2010	Length of respiratory support, PICU LOS, hospital LOS	nCPAP reduction in invasive care, LOV, PICU LOS, hospital LOS, and economic burden.	Initial No severity score applied.
NIV – CMV	Ganu[47]	AUS	2012-04	R	S	PICU	520	5/2.78	2000	2009			LOS PICU	Introduction of NIV halved the LOS in intensive care	
CMV/CPAP	Mansbach[48]	USA	2012-08	P	M	Ward, PICU	2207	24/4	2007	2010			Need for CPAP and/or intubation	Factors were age <2 months, maternal smoking during pregnancy, birth weight <2.5 kg, breathing difficulty began <1 day before admission, apnea, inadequate oral intake, severe retractions, room air oxygen saturation <85%	

Ward = General Pediatric Ward, ED- Emergency department (ward), PICU – pediatric intensive care unit, HDU – high dependency unit;
Center: S=Single, M=Multi; Type: R=Retrospective, P=Prospective; HFNC - heated humidified High-Flow Nasal Cannula, CPAP=continuous positive airway pressure, CMV=mechanical (invasive) ventilation, NIV=non-invasive ventilation, GCS=Glasgow Coma Scale, LOV=Length of ventilation

Table 4: Randomized Controlled Trials, 2010-2018 Investigating Non-Invasive Ventilation

AUTHOR	YEAR	SETTING	AGE	N	INTERVENTION (N)	COMPARATOR (N)	OUTCOME	RESULTS	COMMENTS
Milési[49]	2018	Multicentre (only PICU)	0-6m	286	HFNC - Flow 3 L/Kg/min (144)	HFNC - Flow 2 L/kg/min (142)	Treatment failure within 48h after randomization as defined by 1 of 4 criteria (mWCAS, RR, EDIN, apnoea)	HFNC at a flow of 3 L/Kg/min did not reduce failure rate. Instead, it seemed to increase discomfort and possibly increase length of stay.	One of the inclusion criteria was mWCAS≥3. However, mWCAS is only partially validated. No comparison made to standard care or nCPAP.
Franklin[16]	2018	Multicentre (non PICU)	0-12m	1472	HFNC (739)	Standard Care (733)	Escalation of care (≥3 of 4 clinical criteria)	HFNC had significantly lower escalation rate than standard care	Length of stay was not different between HFNC and standard care. Reason for escalation was often outside the predefined 4 criteria ("clinician's decision").
Milési[13]	2017	Multicentre	0-6m	142	HFNC (71)	nCPAP (71)	Treatment failure within 24h after randomization (includes mWCAS)	nCPAP potentially more efficient than HFNC for initial treatment of AVB	mWCAS only partially validated
Kepreotes[18]	2017	Single center, not blinded, mild disease	0-24m	202	HFNC (101)	LFNO (101)	Time from randomization to last use of oxygen therapy	No significant difference between HFNC and LFNO	HFNC might be useful as rescue therapy to prevent costly PICU treatment
Chidini[50]	2015	Multicentre (3 PICUs), not blinded, mild disease	6-12m	30	CPAP via helmet (17)	CPAP via facial mask (13)	Treatment failure Secondary: CPAP application time, number of patients requiring sedation, complications	CPAP-helmet better tolerated than CPAP-facial mask and requires less sedation.	Small study. Assessment of severity not described
Bueno Campaña[51]	2014	Two hospital wards, moderate disease (RDAI ≥4)	0-6m	75	HFNC (32) + epinephrine	Inhaled hypertonic saline (43) + epinephrine	RACS Secondary: mean comfort score, LOS, PICU admission	No difference with respect to severity and comfort scores, LOS or PICU admission rate.	RDAI/RACS poorly validated
Milési[52]	2013	Single center, PICU, severe disease, randomization for 6 h	0-6m	19	nCPAP (10)	LFNO (9)	mWCAS, inspiratory muscle work measured by esophageal pressure	nCPAP rapidly decreased inspiratory work	Small study. mWCAS not validated for this purpose
Hilliard[53]	2012	Pilot study, hospital, moderate disease	0.3 - 11.3m median 3m	19	HFNC (11)	Head box oxygen (8)	SpO₂ 8h post randomisation	Median SpO₂ was higher in the HFNC group at 8h (100% vs 96%), and 12h (99% vs. 96%), but similar at 24 h.	Small study. SpO₂ difference clinically relevant?
Kim[54]	2011	Single centre, ED, severe disease defined by m-WCAS >3	2-12m	69	Heliox inhalation (34) via HFNC + albuterol + epinephrine	Oxygen inhalation (35) via HFNC + albuterol + epinephrine	mWCAS up to 4hrs or until emergency department discharge secondary: RDAI	Epinephrine delivered by helium-oxygen inhalation, followed by helium-oxygen HFNC improved m-WCAS more than just oxygen	RDAI poorly validated. mWCAS only partially validated

HFNC - Heated humidified high-flow nasal cannula. LFNO - low flow nasal oxygen. EDIN - neonatal pain and discomfort scale
m-WCAS - modified Wood's clinical asthma score. RACS = Respiratory Assessment Change Score. m – month(s)

The sample sizes of seven studies were between 100 and 600 subjects. One study examined more than 2000 subjects, which was the only study with a prospective design and multi-centre design.[48] Four studies compared two time periods (pre–post intervention study). The year 2010 roughly marked the introduction of high flow nasal cannula in clinical practice. Four studies dealt with high flow therapy, examining success and failure rates, and investigating the introduction of a protocol for the general ward. The remaining studies, which mostly analysed datasets collected before the year of 2010, described the use of nasal continuous positive airway pressure (nCPAP) in comparison with conventional mechanical ventilation, predictors of nCPAP, or compared non-invasive ventilation with conventional mechanical ventilation.

It was shown in two reviews how the introduction of non-invasive ventilation (mainly nCPAP) helped to reduce length of stay in paediatric intensive care and therefore hospital length of stay[47,55]. One study compared the need for intubation before and after the introduction of high flow nasal cannula which was reduced from 23% to 9%. The length of stay in paediatric intensive care decreased from six to four days.[44] A retrospective review of emergency department case notes of infants less than twelve months of age found that 28 out of 163 (17%) required nCPAP. The predictors were oxygen requirement within the emergency department, lower oxygen saturation, younger age at presentation, higher respiratory rate, higher heart rate, lower Glasgow Coma Scale score, and younger gestational age than the group that did not require non-invasive ventilation.[45]

A prospective, multi-centre cohort study in 2007-2010 looked at factors associated with ventilation in bronchiolitis patients younger than two years of age. It was found that 163 out of 2207 patients (7.4%) required ventilation (nCPAP or conventional mechanical ventilation). The factors that increased the risk for ventilation were: age less than two months, birth weight less than 2.5 kg, apnoea, inadequate oral intake, severe retractions, and oxygen saturations (SpO_2) less than 85% in room air. Newly identified factors were maternal smoking during pregnancy and rapid respiratory decline with breathing difficulty commencing less than one day before admission.[48]

Another study investigated variables that predicted high flow nasal cannula failure. There was no correlation between high flow nasal cannula failure and history of prematurity or patient's age. Non-responders to high flow nasal cannula had higher partial pressure of carbon dioxide (pCO_2) and less tachypnoea at the initiation of high flow nasal cannula than responders. In addition, non-responders had no change in respiratory rate during high flow nasal cannula and had higher paediatric risk of mortality scores in the first 24 hrs.[43]

Most recently, two articles described a 24-months period before and after the introduction of a protocol for high flow nasal cannula for the general paediatric ward. An inpatient classification system (All Patients Refined Diagnosis Related Groups = APR-DRG) was used as a clinical severity score. The first analysis showed reduced length of stay for bronchiolitis patients after

introducing high flow nasal cannula to general wards. However, the second analysis revealed no effect for hospitalized patients on length of stay, length of stay in paediatric intensive care, transfer rate to paediatric intensive care, conventional mechanical ventilation rate, or 30-day re-admission rate. The study material provided data to calculate length of stay in days only. The use of high flow nasal cannula increased from 24% to 35%. It was concluded, that the increased availability in the ward might have resulted in overuse of high flow nasal cannula.[17,42]

This review found nine RCTs in relation to non-invasive ventilation and acute viral bronchiolitis (Table 4, page 23). End points and outcomes of the studies were treatment failure, length of oxygen therapy, oxygen saturation, and clinical severity scores. The sample size was under 70 subjects for four studies. There were four studies that had a multi-centre design. Six RCTs used a clinical severity score as an outcome variable or as part of the outcome (e.g. treatment failure).

A small study comparing nCPAP with low flow nasal oxygen used the modified Wood's Clinical Asthma Score (mWCAS) and oesophageal pressures to demonstrate how nCPAP rapidly decreased inspiratory work.[52] It was pointed out by other researchers that interpretation of these results are problematic because mWCAS is not validated for this purpose.[14] A small pilot study comparing high flow nasal cannula with head box oxygen used pulse oximetry (SpO2) eight hours after randomization as outcome measure. There was a statistically significant difference between the two groups at 8 hours and 12 hours, but not at 24 hours.[53]

The comparison of two types of nCPAP application demonstrated that helmet nCPAP was better tolerated than facial nCPAP and required less sedation. However, it is unclear how disease severity was assessed and whether only infants with mild disease were treated.[50] Another study compared high flow nasal oxygen either given together with Helium (Heliox) or without which was applied in combination with albuterol and epinephrine. Outcome measures were two clinical severity scores, mWCAS and RDAI, which were not validated for this purpose.[54] Another study found no difference with regards to change in clinical severity score and other parameters when comparing the application of high flow nasal cannula with inhalation of hypertonic saline. The use of a poorly validated score (RDAI/RCAS) makes interpretation of these results difficult.[51]

Two larger trials compared high flow nasal cannula with other treatment options. One study compared high flow nasal cannula with low flow nasal oxygen. The primary outcome was time from randomization to last use of oxygen therapy. It was concluded that high flow nasal cannula did not significantly reduce time on oxygen compared with standard therapy.[18] The second study compared high flow nasal cannula with nCPAP for the initial treatment of acute viral bronchiolitis. Primary outcome was treatment failure within 24 hour of randomization which was defined as one or more criteria met out of four: increase in severity score, increase in respiratory rate, increase in

discomfort score (EDIN) and occurrence of apnoea. The authors concluded that "*nCPAP may be more efficient than HFNC for initial respiratory support*".[13] Another study by the same first author, found no difference between application of high flow nasal cannula with a flow of 3 L/kg/min and 2 L/kg/min in relation to treatment failure. Again, mWCAS, an only partially validated respiratory severity score, was used as an entry criterion and was one of the four criteria for failure. Some of the limitations were that the study was not blinded and that nCPAP was not part of the study as an alternative treatment.[49]

Another large trial from Australia and New Zealand randomly assigned bronchiolitis patients (n=1472) to either standard therapy or high flow therapy (PARIS trial).[16] The main result of the study was that high flow therapy had significantly lower escalation rate than standard care. The escalation rate in the high-flow group was 12% versus 23% in the standard-therapy group. Escalation of care was defined as not meeting three out of four clinical criteria (heart rate, respiratory rate, oxygen requirement, PEWS alert triggered) or if the clinician decided to escalate for other reasons. The clinician's decision to escalate a case, without the patient meeting at least three criteria of the predefined four criteria, occurred in 34%. This suggests insufficient criteria of failure for describing the patient's state of illness. Analysis of secondary outcomes revealed no differences between high flow therapy and standard care in length of hospital stay, length of stay in intensive care, or the duration of oxygen therapy. Because severity of disease was not an inclusion criterion, unnecessary use of high flow therapy in patients with mild disease could have been the reason for the lack of difference in secondary outcomes.

2.3 Discussion

2.3.1 Clinical Severity Score

There is no gold standard for a clinical severity score.[34] Only limited data is available on validation, despite the fact that the COSMIN initiative provided guidance on how to assess reliability, validity, responsiveness, and interpretability of a health measurement instrument.[56,57] The AAP Guideline commented on clinical severity score, *"none has achieved widespread acceptance and few have demonstrated any predictive validity, likely because of the substantial temporal variability in physical findings in infants with bronchiolitis"*.[6]

Currently, the modified Wood's clinical asthma score (mWCAS) appears to be the most often used score for the severity assessment for patients with acute viral bronchiolitis despite not being fully validated.[34,58,59], To the author's knowledge, van Miert et al. were the first to develop the most comprehensive and most validated score for acute viral bronchiolitis (LIBSS-PRO). However, it still lacked validation for responsiveness to change and cross-cultural validation identifying differences in language and cultural practices.[41]

A fully validated and internationally accepted clinical severity score should be an invaluable tool for defining admission and discharge criteria. It would help to define treatment failure, help to recognize short and long-term effects of therapeutic interventions, and help to reduce heterogeneity of research studies, thus facilitating meta-analyses. Therefore, highlighting the importance of assessing the severity of the disease. International research has not been able to fully validate a clinical severity score so far. The lack of a gold standard, with many items being subjective, especially when auscultation skills are required, makes validation difficult. In addition, one should consider correcting heart rate and respiratory rate for body temperature.

It remains to be seen whether existing EHR systems can provide the information required to estimate clinical severity, or whether more precise and robust data input is required. Machine-learning (ML) techniques have been most successful in predictive analytics, as soon as the "right data" became available.[21,60] Increasingly, genetic data is used to customize medical treatment based on individual characteristics of patients and their response to treatment. New techniques of unsupervised feature learning ("Deep Learning") might help to discover new phenotypes at a much larger scale.[61]

2.3.2   Modes of Non-Invasive Ventilation (NIV)

Empirical evidence has shown that the preferred mode of respiratory support is non-invasive ventilation. Nowadays, conventional mechanical ventilation is reserved for severe acute viral bronchiolitis and high-risk cases with co-morbidities. High flow nasal cannula has gained wide popularity despite the lack of clearly established evidence, mainly because of its simplicity and excellent tolerance.

In a recent review that listed empirical data from observational studies, Sinha and colleagues commented, "*At best, observational studies such as these provide some indication, rather than direct evidence…*". The research focus has gradually moved away from conventional mechanical ventilation to non-invasive ventilation. There is ongoing debate about the mode and timing of non-invasive ventilation. The results of the large PARIS trial have demonstrated that high flow therapy is an important rescue therapy for failed low flow therapy. However, the trial did not find improved length of stay for patients treated with high flow therapy. This was also found by another Australian study.[18] Milési and colleagues concluded in a recent RCT that nCPAP was potentially more efficient than high flow therapy. At present, it is not clear whether escalation of care for patients treated with high flow therapy should include nCPAP in addition to conventional mechanical ventilation.

2.3.3   Limitations of this Review

This review has limitations in relation to the correct identification of relevant research articles. Only articles written in English were included. Despite using multiple research strategies

finding articles relevant to acute viral bronchiolitis in infants was challenging. Many of the articles in this review did not have MeSH terms linked to them, which made them difficult to find.

Another systematic limitation is the lack of an internationally accepted definition of Acute Viral Bronchiolitis. There is international agreement that "*clinicians should diagnose bronchiolitis and assess disease severity on the basis of history and physical examination*".[6] However, there is no universally accepted definition of acute viral bronchiolitis. The 2016 AAP guidelines state "*Most clinicians recognize bronchiolitis as a constellation of clinical signs and symptoms occurring in children <u>younger than 2 years</u>, including a viral upper respiratory tract prodrome followed by increased respiratory effort and <u>wheezing</u>.*"

The UK, Australia and parts of Europe, define the term "acute bronchiolitis" as an apparent viral infection of an infant <u>younger than one year</u> who develops signs of lower respiratory tract disease with airflow obstruction manifested as increased work of breathing, hyperinflation of the chest and widespread crackles; they <u>may or may not intermittently exhibit wheeze</u>.[62,63] Studies based on the U.S. definition tend to represent slightly older infants and toddlers, many of whom would have a wheezy bronchitis and some of whom may have asthma.

2.4 Key Considerations for this Study

So far, there is no internationally accepted and fully validated clinical severity score. A fully validated score could assist with treatment decisions, measuring outcome, predicting outcome, and evaluation of new therapies. EHR data entered may provide enough material to enable validation on a large scale.

Non-invasive respiratory support for acute viral bronchiolitis is now the standard for mild to moderate disease. The recent RCTs have provided further knowledge for the optimization of NIV therapy. High flow therapy is now recognized as an important rescue therapy. So far, it is not known whether nCPAP should be used for escalation of care. It is also not known which mode of non-invasive ventilation should be used at what level of disease severity.

3 Data

The following four subsections (3.1-3.3) briefly explain the methodology on how to obtain a dataset of this size (e.g. 14 million event records). Finally, subsection 3.4 describes the resulting dataset. Basic descriptive statistics and graphs are used to describe relevant characteristics.

3.1 Data Acquisition

The CERNER Millenium Database is a complex relational database management system with multiple tables. Information services at CHLA granted the author direct, read-only access to the CERNER database via the *Discern Query Builder* which provided comprehensive information about the data tables and columns (Fig. 1).
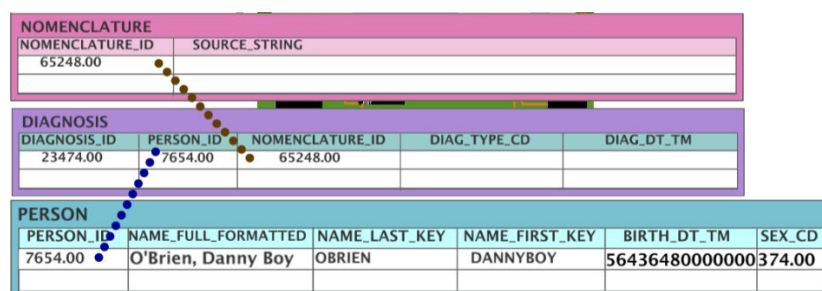
Fig. 1: Example of three related tables in CERNER Millenium Database

The *Discern Query Builder* assisted with creating complex SELECT statements which were written in the Cerner Command Language (CCL). CCL resembles in parts the Structured Query Language (SQL). The graphical user interface of the *Discern Query Builder* allowed instant retrieval of electronic health record data. The main tables queried for this study were: *nomenclature, diagnosis, person, encounter, clinical_event*. The results were exported to comma-separated values (CSV) files which were stored in a folder located on the hospital network.

At this stage, the only identifiable patient information was the date of birth which was subsequently converted into minutes relative to a randomly selected date. Every data field that contained date and time information was converted into minutes since date of birth, which was also expressed in minutes, thus preserving the temporal relationship between data points.

Data extraction was performed in five major steps (see appendix for examples of CCL code):

1. Extract diagnoses (466.1*, J21.*) and find accompanying patients
2. Extract Patient details (DOB, Sex, Prematurity, birth weight)
3. Extract encounters
4. Extract all other diagnoses per patient and encounter
5. Extract all clinical events for each patient and encounter

As a result, the raw dataset contained 7,043 patients comprising 8,317 encounters, 23,691 diagnoses and 14,438,356 clinical events. After inspecting the raw dataset, the author decided against downloading the accompanying clinical free text. The downloaded structured text data was

considered sufficient to answer the research question of this study. In addition, it was felt that current natural language processing was still prone to error and misinterpretation, and therefore time consuming and beyond the scope of this study.

Patient Cohort

The CERNER Millenium database was screened for all patients with the diagnosis "Acute Viral Bronchiolitis" represented by the ICD codes 466.* and J21.*. As per the institutional review board, only patients admitted between January 2008 and March 2017 were included.

```
ICD 10
▶ J21 Acute bronchiolitis
      ▶ J21.0 … due to respiratory syncytial virus
      ▶ J21.1 … due to human metapneumovirus
      ▶ J21.8 … due to other specified organisms
      ▶ J21.9 … unspecified
ICD 9
      ▶ 466.1 Acute bronchiolitis
            ▶ 466.11 … due to respiratory syncytial virus (RSV)
            ▶ 466.19 … due to other infectious organisms
```
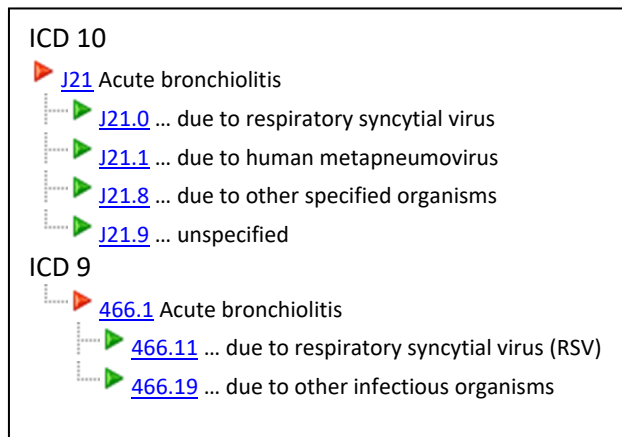
Fig. 2: ICD Codes Used for Identification of Patient Cohort

3.2 Data Preparation

De-Identification

The raw dataset contained protected health information: date of birth, medical record number (MRN), date and time of each clinical event, date and time of admission, and date and time of discharge from hospital. The raw dataset was stored on a local hospital network drive and accessed on a PC in a protected environment. The author developed a computer program and used SQL scripts to apply the following, mandatory steps which ensured a fully de-identified, HIPAA-compliant dataset whilst preserving the temporal relationship of the data.

1.   Remove MRN and create random patient IDs and random case numbers

2.   Remove any information related to rare diseases or conditions

3.   Remove any personal information from clinical event table

4.   Calculate day of the week, minute of the day and month for each clinical event

5.   Calculate time difference between date of birth and clinical event

6.   Convert date of birth into minutes

Errors of User Entry

The datafiles extracted from the CERNER database stored the values and text-input of clinical events as alpha-numerical values. The EHR system (*PowerChart*) allowed the user to enter text data in numerical fields that should only contain plain integer or floating-point values. As a

result, approximately 3-5% of the data were invalid. In most cases, the author's scripts or sometimes manual analysis replaced the invalid number by the average value of the preceding and following value found in the database. In some cases, visual inspection of the patient chart and/or data tables revealed correct values be inspecting adjacent data fields. For example, the values of heart rate (HR) and respiratory rate (RR) needed to be swapped because the user had obviously entered heart rate and respiratory rate in wrong order, e.g. HR 23bpm and RR 145bpm instead of HR 23bpm and RR 145bpm.

Fraction of Inspired Oxygen (FiO$_2$)

Any application of supplemental oxygen delivers an oxygen concentration that is above the normal oxygen concentration in room air of 21%. The fraction of inspired oxygen (FiO$_2$) describes the amount. Usually, the normal range for FiO$_2$ is 0.21-1.0 or when expressed in percent, 21% to 100%. In rare instances, the given FiO$_2$ is below 0.21 (or 21%), usually applied to cardiac patients to reduce pulmonary blood flow. However, this is not anticipated in patients with increased oxygen demand, i.e. acute viral bronchiolitis. The following invalid numbers, including text entry, were found in the extracted datafile: *0, 0.000, 0.125, 10, 101, 102, 103, 14.000, 17, 18, 2, 20, 211, 213, 3, 4, "40% hi flow", 5, 6, 7.*

Pulse Oximetry

Pulse oximetry is a non-invasive method for monitoring a patient's oxygen saturation (SpO$_2$) in the blood stream. Normal readings range from 95 to 100 percent in a healthy person. The following invalid numbers were found in the extracted datafile: *0, 0.250, 0.5, 10, 11, 12, 14, 15, 16, 17, 19, 2, 4, 5, 6, 7.*

Heart Rate (HR)

The following invalid or highly unlikely numbers were found in the extracted datafile: *0, 1, 10, 10236, 11, 1112, 1470, 1541, 16243, 1665.*

Respiratory Rate (RR)

The respiratory rate was often entered as greater than 250 bpm.

Flow rate (L/min)

The flow rate was often reported as greater than 100 L/min which is technically not possible. To correct these invalid numbers, the same algorithms, as described above, were applied.

Body Weight

The body weight at the time of admission to hospital often varied significantly from other weight measurements during the stay. To obtain a useful reference weight for the final analysis, the author developed a correction algorithm which determined the median weight of all available measurements per patient and encounter.

Respiratory Support Therapy

The clinical event *Type of Oxygen Administration* described how oxygen was applied. The most common types were nasal cannula, conventional mechanical ventilation, high flow nasal cannula and nasal continuous positive airway pressure (nCPAP).

Depending on the location of the patient, *Type of Oxygen Administration* was usually accompanied either in critical care by FiO_2 and oxygen/air flow, or only by flow when the patient's location was the general ward. The fact that high flow nasal cannula was only applied in critical care lead to a high correlation between high flow nasal cannula and FiO_2. Therefore, the covariable FiO_2 was eliminated from the final data analysis. Estimation of FiO_2 by using empirical oxygen values for flow rates were not deemed accurate enough.

The ending date and time of respiratory therapy was in many cases not recorded. Manual correction was applied by examining factors like oxygen flow documented as 0.0 L/min, patient discharged from hospital or, in the case of high flow, discharged from critical care. In addition, the clinical event *Room Air="Yes"* was used as an indicator of discontinuation of respiratory support therapy. Each episode of non-invasive ventilation was manually selected and marked for analysis.

Length of Stay

The CERNER EHR system recorded the length of every patient encounter by documenting the exact date and time of the patient's admission to hospital and discharge. Respiratory support therapy was often commenced in the emergency department which therefore occurred prior to admission to hospital. Consequently, date and time stamp of the first event recorded in the emergency department was used to calculate the total length of stay, expressed in minutes. This variable was used as the primary outcome variable in the data analysis.

Some researchers have made the distinction between actual discharge from hospital and the patient's readiness to leave the hospital because non-medical reasons, e.g. social factors and/or availability of transport, can sometimes prolong the length of stay. The attribute, "Patient ready for discharge", usually means that the patient is free from symptoms and has not received any respiratory support therapy for the last 6-12 hours.[64] This study did not attempt to adjust the length of stay because of insufficient data.

3.3 Data Conversion

3.3.1 Calculating Centiles for Heart Rate, Respiratory Rate and Blood Pressure

Instead of basing age dependent variables on absolute values, the author used recently published centile curves for heart rate, respiratory rate, systolic and diastolic blood pressure to build a score as shown in Table 5.

Table 5: Score Values Based on Centiles

Centile	Score Value
≤ 90th	0
>90th and ≤ 95th	1
>95th and ≤ 99th	2
>99th	3

The underlying data were extracted from Bonafide et al.[65] who developed and validated heart and respiratory rate percentile curves for hospitalized children. The dataset for blood pressure was recently published by Eytan et al.[66]

3.3.2   Prematurity - Gestational Age

The clinical event table contained the event type *Gestational Age* which documented the age at the time of birth.

3.3.3   Body weight at admission (z score)

The data of the Center for Disease Control and Prevention (CDC, Atlanta, USA) were used to calculate the z score of the body weight. The dataset provided by Fenton et al. was used to calculate the z score of the body weight for premature babies.[67]

3.4 Results

3.4.1   Patient Characteristics

The final dataset contained all patients under the age of one year who presented with acute viral bronchiolitis to the Children's Hospital of Los Angeles between 01/2008 and 03/2017 and were treated with non-invasive ventilation.

Fig. 3 depicts the selection process from the initial figure of 8,317 encounters with acute viral bronchiolitis as part of the problem list to the final figure of 592 encounters for 579 patients.

Only cases with respiratory support therapy, either standard nasal cannula, high flow nasal cannula or continuous mechanical ventilation, entered the final analysis.

Table 6 describes the standard and high flow group in detail. The baseline demographic and physiological characteristics were similar between both groups.

Screen for Acute Viral Bronchiolitis

8,317 encounters for 7,048 patients

Find clinical event "Type of Oxygen Administration"

2,075 encounters for 1,511 patients

Limit Dataset to
- 01/2008 and 03/2017 (as per IRB)
- age at admission 0 – 365 days
- Length of Stay 2-21 days

827 encounters for 794 patients

Apply Further Exclusion Criteria

235 encounters excluded

Cnt	Reason for Exclusion
101	Severe lung disease
49	Outlier - tLOS >15,000 min (>10.42 days)
27	Resp. support other than NC or HFNC
26	Severe heart disease
10	Severe upper airway disease
6	Severe systemic disease
5	Insufficient data
5	Whooping cough
4	Rare genetic disease
2	Death

Final Dataset

592 encounters for 579 patients

Count	Description
13	Re-admissions
258	Simple Cases
334	Complex Cases
588	Discharge Home
470	Standard resp. therapy (NC)
122	High Flow Nasal Cannula (HFNC)
18	HFNC escalated to CMV
35	NC escalated to HFNC

Fig. 3: Patient Selection Process

Table 6: Patient Characteristics

Variable	Standard Group	High Flow Group
Total	**470**	**122**
Treatment escalated	35 (7.4%)	18 (14.8%)
Simple Cases (Bronchiolitis only diagnosis)	209 (44.5%)	49 (40.2%)
Simple Cases - Treatment escalated	16 (3.4%)	5 (4.1%)
Complex Cases - Treatment escalated	19 (4.0%)	13 (10.7%)
Z score of body weight at admission	-0.27 ±1.49	-0.19 ±1.47
Age at Admission [days]	142 ±105	134 ±106
total Length of Stay (tLOS) [days]	4.4 ±1.8	5 ±1.9
Female Sex no.	204 (43.4%)	48 (39.3%)
Prematurity no.	26 (5.5%)	7 (5.7%)
Viral Cause		
Number of cases tested for viral cause	286 (60.9%)	108 (88.5%)
RSV no.	225 (47.9%)	74 (60.7%)
virusRhinoEntero no.	16 (3.4%)	15 (12.3%)
virusMetapneumo no.	23 (4.9%)	15 (12.2%)
virusParainf3 no.	13 (2.8%)	
virusAdeno no.	9 (1.9%)	4 (3.3%)
TOP 10 Diagnostic Codes (ICD 10)		
Other acute lower respiratory infections (J20 - J229)	470 (100%)	122 (100%)
Symptoms and signs involving the circulatory and respiratory systems (R00 - R099)	350 (74%)	106 (87%)
Influenza and pneumonia (J09 - J189)	147 (31%)	34 (28%)
Bacterial and viral infectious agents (B95 - B979)	120 (26%)	26 (21%)
Other diseases of the respiratory system (J96 - J999)	117 (25%)	68 (56%)
Other respiratory diseases principally affecting the interstitium (J80 - J849)	101 (21%)	25 (20%)
Metabolic disorders (E70 - E889)	94 (20%)	18 (15%)
General symptoms and signs (R50 - R699)	92 (20%)	15 (12%)
Congenital malformations of the circulatory system (Q20 - Q289)	67 (14%)	12 (10%)
Respiratory and cardiovascular disorders specific to the perinatal period (P19 - P299)	56 (12%)	(null)

3.4.2   Distribution of Hospital Length of Stay

The distribution of the total length of Stay expressed in minutes demonstrated a right (positive) skew (Fig. 4).
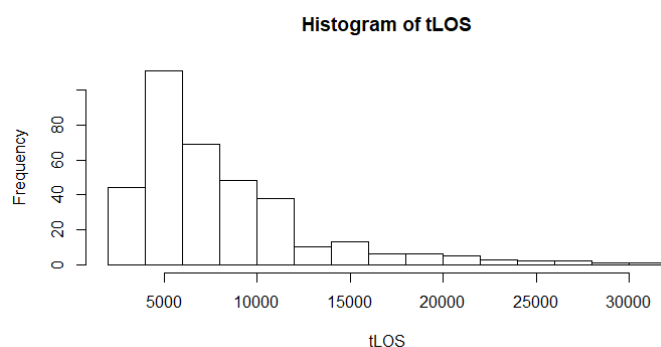
Fig. 4: Histogram of total Length of Stay including outliers

The unadjusted dataset contained 54 outliers with a length of stay greater than 15,000 minutes (>10.42 days). These patients were younger than 100 days, 87% were complex patients

who in many cases required prolonged mechanical ventilation. In view of the average length of stay of 5 days, the likelihood of other non-related to acute viral bronchiolitis or non-medical factors were considered and an upper limit of 15,000 minutes for total length of stay applied.

The resulting distribution of length of stay is shown in Fig. 5 which depicts the two treatment groups (standard care and high flow). There were nearly four times as many cases in the control group than in the treatment group. Most cases in the standard group had a length of stay less than 5 days (<7200 min).
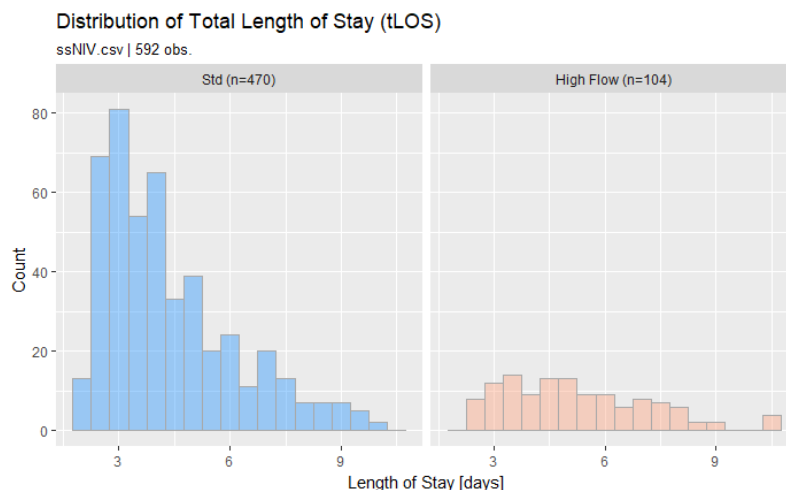
Fig. 5: Distribution of total Length of Stay

3.4.3   Age Distribution at Admission.

The age at the time of admission is depicted in the histogram below (Fig. 6). The age was recorded in days and, if necessary, corrected for gestational age.
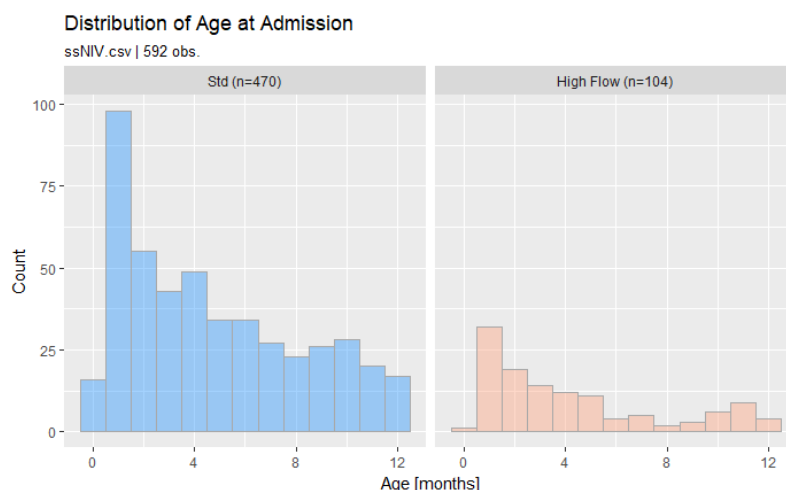
Fig. 6: Distribution of Age at Admission

The largest group of patients was infants less than three months old (

Table 7) followed by 4-6 months old children. These two age groups represented nearly 70% of all patients.

Table 7: Age distribution Divided by Quarter

NIV	CaseType	Cnt	Male	Prem	Count Age 1-3 mon	Count Age 4-6 mon	Count Age 7-9 mon	Count Age 10-12 mon
NC	Simple	261	145	19	91	75	46	49
NC	Complex	209	121	7	99	53	34	23
HFNC	Simple	41	27	1	12	16	5	8
HFNC	Complex	28	18	2	13	3	7	5
NC escalated	Simple	19	7	1	10	7		2
NC escalated	Complex	16	10	1	7	4	2	3
HFNC escalated	Simple	13	8	2	10	1		2
HFNC escalated	Complex	5	4	0	5			
All Cases		**592**	**340**	**33**	**247 (42%)**	**159 (27%)**	**94 (16%)**	**92 (16%)**

NC – Nasal cannula (=standard group), HFNC – high flow nasal cannula

3.4.4   Co-morbidities and Risk Factors

Co-morbidities

The diagnostic codes used in this study were the codes of the international classification of diseases, ICD-9 and ICD-10. For convenience, ICD-9 codes were converted into ICD-10 codes because it was not an objective of this study to examine the ICD codes in detail. The top 10 of the main ICD groups are listed in Table 6 on page 36. For each case, manual inspection of the available clinical data was performed to apply further exclusion criteria. As listed in Fig. 3 on page 35, the five most common reasons for exclusion were severe lung disease, LOS outliers of greater than 15,000 minutes (>10.42 days), respiratory support other than standard therapy or high flow therapy, sever heart disease, and severe upper airway disease.

The main difficulty for interpretation of the codes was the lack of grading. For instance, in many cases lack of detailed diagnostic data made it difficult to determine the clinical significance of a code such as *Q21.0* (*Ventricular Septal Defect*). The clinical spectrum of a ventricular septal defect can vary from congestive heart failure with failure to thrive to a benign muscular VSD without any clinical significance. Because of limited diagnostic data in this dataset, it was not attempted to introduce a stratification based on ICD codes except for a binary decision, simple or complex. Cases with acute viral bronchiolitis as the principal diagnosis and no other significant diagnoses were labelled as "simple".

Risk Factors

Table 8 lists the risk factors documented at admission to hospital. It only shows the "Yes" answers. "No" or "Unknown" are not shown. The first field, "Smoker in House", reported a percentage of 6.8% which was most likely under-reported. The remaining fields were useful in a clinical context, serving as warning flags for the treating team, however, due to their general nature they were not included in the risk analysis.

Table 8: Risk Factors recorded in Cerner Millenium Database

	Total	Smoker in House	Previous Medical Problems	Previous Surgery	Cardiovascular Problems
NC (0)	470 obs.	34 (7.2%)	159 (33.8%)	55 (11.7%)	44 (9.4%)
HFNC (1)	69 obs.	2 (2.9%)	16 (23.2%)	5 (7.2%)	6 (8.7%)
NC escalated (2)	35 obs.	3 (8.6%)	11 (31.4%)	2 (5.7%)	1 (2.9%)
HFNC escalated (3)	18 obs.	1 (5.6%)	7 (38.9%)	0 (0.0%)	1 (5.6%)
All Cases	**592 obs.**	**40 (6.8%)**	**193 (32.6%)**	**62 (10.5%)**	**52 (8.8%)**

NC – Nasal cannula (=standard group), HFNC – high flow nasal cannula

Viral Cause

The Los Angeles dataset provided information for five different types of viral causes. As per Table 6 (page 36), 60.9% (n=286) of the cases in the standard group were laboratory tested for a viral cause versus 88.5% (n=108) in the high flow group. Respiratory syncytial virus (RSV) was the most common cause for bronchiolitis (50.3%). Metapneumo virus was the second most common viral agent. Other causes with lower frequency are listed in Table 9. In approximately one third of the cases, viral diagnostics did not yield a positive result.

Table 9: Viral Cause of Bronchiolitis

Virus	Count (%) N=592	Standard Group N=470	High Flow Group N=122
RSV	299 (50.3%)	225 (47.9%)	74 (60.7%)
Rhino-Entero Virus	31 (5.2%)	16 (3.4%)	15 (12.3%)
Metapneumo Virus	38 (6.4%)	23 (4.9%)	15 (12.3%)
Parainf. 3 Virus	13 (2.2%)	13 (2.8%)	
Adeno Virus	13 (2.2%)	9 (1.9%)	4 (3.3%)
No virus detected	199 (34%)	184 (39.1%)	14 (11.5%)

3.4.5   Gender Distribution and Prematurity

The ratio between male and female patients was 1.35 with a percentage of 57% for male patients (340 out of 592). The distribution of gestational age is shown in Fig. 7. There were 33 (5.5%) premature patients recorded in the database which appears a small number given the fact that 57 patients (9.6%) had a z score of the body weight below -2.0. which equals to below 2nd centile. This consideration was supported by the fact that gestational age, either normal or below 37 weeks representing prematurity, was only reported in 97 cases (16%). Because gestational age was not well captured in the dataset and was often missing, it was not included in the analysis.
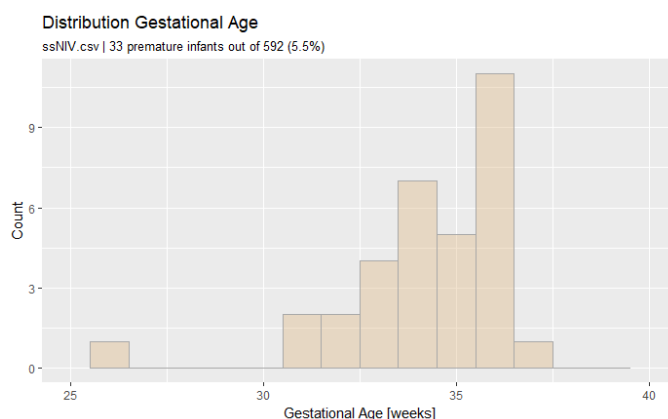
Fig. 7: Distribution of Gestational Age

3.4.6   Distribution of the z score of body weight at admission

The z score (synonym standard score) is the signed number of standard deviations by which the value of an observation or data point differs from the mean value of what is being observed or measured. The z score of the body weight was calculated by using the LMS method provided by the CDC.[68] The z score is based on age and gender. In this study, the age was corrected for the gestational age. For example, the age of a child born at a gestational age of 32 weeks and a chronological age of three months was corrected by minus eight weeks. The resulting age of one month was then used for the z score calculation.

The advantage of the z score is that it describes the distance between the individual weight and the average weight of comparable children. In a normally distributed population 95% of the measurements lie between the mean plus and minus two standard deviations ($\bar{x} \pm 2SD$). Therefore, any z score below minus two or above plus two standard deviations represents an abnormal value.

Because of multicollinearity between body weight and age one must decide to use either body weight or age for statistical analysis. However, the z score of the body weight provides information about the physical development of the child thus converting body weight into a meaningful covariate.
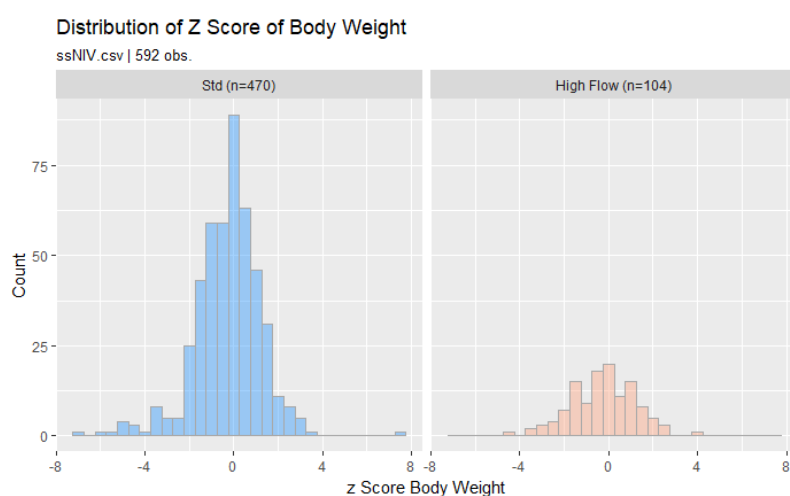
Fig. 8: Distribution of z score of body weight

Fig. 8 depicts a nearly normal distribution of the z score of the body weight for the standard and the high flow group. Most patients had a normal body weight as expressed by a z score between -2.0 and +2.0 (n=508, 86%). The number of patients with a z score below -2.0 was 57 (9.6%) and above 2.0 was 27 (4.6%).

3.4.7   Re-admitted Patients

Any readmission within seven days was a reason for exclusion from the dataset. Table 10 shows the details of the patients who were re-admitted; usually after 30 days of the previous discharge. There were two patients with a 14-day and 27-day time interval between discharge and re-admission.

Table 10: Re-admitted Patients

	Patient No.	First Case No.	LOS [days/hrs]	- - -	Second Case No.	LOS [days/hrs]	Time since last discharge [min]	Time Period since last discharge [days/hrs]
1	168	461	8d 10h	- - -	620	8d 21h	21387	14d 20h
2	705	211	3d 2h	- - -	849	6d 20h	39372	27d 8h
3	640	703	2d 14h	- - -	704	2d 15h	43970	30d 12h
4	289	316	6d 13h	- - -	157	3d 2h	49085	34d 2h
5	667	405	3d 8h	- - -	739	2d 14h	51062	35d 11h
6	497	900	3d 16h	- - -	367	4d 5h	60173	41d 18h
7	485	811	7d 5h	- - -	515	3d 22h	112152	77d 21h
8	751	706	2d 19h	- - -	828	2d 16h	148621	103d 5h
9	492	886	4d 18h	- - -	516	2d 21h	171250	118d 22h
10	464	348	2d 3h	- - -	665	6d 6h	185100	128d 13h
11	194	117	2d 6h	- - -	299	2d 8h	310335	215d 12h
12	252	133	5d 22h	- - -	574	2d 21h	345576	239d 23h
13	602	518	4d 2h	- - -	800	4d 18h	410940	285d 9h

3.5 Conclusion

The above detailed description of data processing demonstrates the current state of EHR systems which are optimized for clinical documentation and billing but not research. Data validation, cleaning and conversion still represents the most important and most time-consuming part of a clinical data analysis project.[22,23]

The final dataset contained enough patients for each treatment group. Although the control (standard) group was nearly four times larger than the treatment group (high flow) the balance with regards to important patient characteristics between both groups was satisfactory. The only major differences in relation to ICD groups J96-J999 and P19-P229 were noted.

4   Severity Score for Respiratory Distress

As discussed in the *Literature Review* (chapter 2, p. 18), there are many severity scores for respiratory distress, however, none of the listed scores are fully validated and internationally accepted. Therefore, this study could not simply rely on a gold standard and extract known attributes of a fully validated score from the CERNER database. In addition, due to the retrospective nature of this study, many items were not systematically obtained thus analysis of individual items would not have been representative in many cases. Manual inspection of the cases revealed that many items were not equally recorded. Many clinical parameters and observations were not recorded frequently enough, especially in simple cases.

In the search for suitable parameters stored in the EHR system, the Los Angeles dataset was found to contain 22 different data fields that were considered suitable for describing severity of respiratory disease. These data fields comprised eleven numerical parameters (heart rate, respiratory rate, non-invasive blood pressure, pulse oximetry, arterial pH, arterial pO_2, and capillary pH, pO_2 and pCO_2) and eleven text parameters (affect/mood of patient, level of consciousness, skin colour, room air, nasal flaring, respiration type, respiratory PEWS, respiratory effort, respiratory retraction, upper airway sounds, breath sounds). The author developed software that applied a standardized algorithm to process each parameter (described in the following section) and convert it into a score.

Table 2 (p. 20) lists the most commonly used clinical severity scores to describe bronchiolitis and dyspnoea. The following table (Table 11) shows that this study added six more items to the score. At this stage, covariates that acted like risk factors, e.g. age, pre-existing conditions and co-morbidities, were not included in the score. The variables *Feeding*, *Urine Output*, *Capillary Refill Time*, and *Dehydration* were excluded because of limited importance with regards to treatment with non-invasive ventilation: They only play a role at the time of hospital admission as part of the very first assessment because as soon as the child is admitted to hospital the routine care ensures normal fluid input and therefore quick normalisation of the other dependent variables. There were no data available for the items *General Condition* and *Duration of Symptoms*.

Table 11: Most Commonly Used Score Items

Score Item	Count out of 15	Used in this study	Comment	Manually assessed
Wheeze	12	✔	Breath Sounds, Upper Airway Sounds	✔
Retractions	12	✔	Respiratory Retraction	✔
Respiratory Rate	11	✔		
Heart Rate	6	✔		
Breath Sounds	5	✔	Breath Sounds	✔
Dyspnoea (WOB)	4	✔	Respiratory Effort, Nasal Flaring,	✔
SpO2	4	✔	Pulse Oximetry	
General Condition	3			✔
Oxygen Need	3	✔	Room Air	
Appearance	2	✔	Affect/Mood of Patient, Level of Consciousness	✔
Feeding	2			
Age	2		Was used as covariable	
Duration of Symptoms	1			
Apnoea	1	✔	Respiration Type	
Cyanosis	1	✔	Skin Colour	
Urine Output	1			
Capillary Refill Time	1			
Dehydration (WHO)	1			
Blood Pressure (sys/diast)		✔		
Temperature		✔		
Arterial Blood Gas – pH		✔		
Arterial Blood Gas – pO$_2$		✔		
Capillary Blood Gas – pH		✔		
Capillary Blood Gas – pCO$_2$		✔		
Capillary Blood Gas – pO$_2$		✔		
Affect/Mood Patient		✔		✔

4.1 Development of a Severity Score

Out of the total of 8,277 different types of events, 22 were selected for assessment of severity of illness. This selection contained eleven clinical events of a numerical data type and eleven of a categorical data type. The following table shows the top ten events (Table 12). The full list of clinical events can be found in the appendix (Table 34, page 123).

Table 12: Top Ten Clinical Events Used for the Severity Score

Name of Data Field	Units	Data Type	Count	Percentage*
Pulse Oximetry	%	NUM	42258	9.28%
Heart Rate	bpm	NUM	38252	8.40%
Respiratory Rate	bpm	NUM	37999	8.34%
Level of Consciousness		TXT	24304	5.34%
Respiration Type		TXT	22411	4.92%
Diastolic Blood Pressure	mm HG	NUM	22032	4.84%
Systolic Blood Pressure	mm HG	NUM	22029	4.84%
Respiratory Effort		TXT	21888	4.81%
Skin Color		TXT	21620	4.75%

*Count/455,381 x 100

The next two paragraphs provide a comprehensive list of the clinical events used in the final analysis. Each event underwent a normalisation process which involved age-specific centiles and conversion to ordinal scales.

4.1.1 Numerical Event Type (Vital Parameters)

The vital parameters used in this study were fraction of inspired oxygen (FiO$_2$), peripheral oxygen saturation (SpO2), heart rate (HR), respiratory rate (RR), non-invasive systolic and diastolic blood pressure, body temperature, arterial pH (ABG_pH), partial pressure of arterial oxygen (ABG_pO2), capillary pH (CBG_pH), partial pressure of capillary oxygen (CBG_pO2), and partial pressure of capillary carbon dioxide (ABG_pCO2).

Because the scale and reference range of vital parameters vary greatly the raw values were translated into clinically relevant scores. The normal range of heart rate, respiratory rate and blood pressure is age-dependent, in 33 cases, it was corrected for gestational age. This study used centile curves for heart rate, respiratory rate and blood pressure developed in hospitalized children. This approach helped not to overestimate the influence of heart rate, respiratory rate and blood pressure.[65,66] The resulting centiles and the absolute values of the remaining parameters were translated into scores by applying clinically relevant ranges as depicted in table (Table 13).

A computer algorithm used a *"greater than"* and *"less or equal"* rule to choose the corresponding score. The score of zero represented a normal physiologic finding. The scores between one and three mirrored common clinical practice to describe severity in ascending order: "mild", "moderate", and "severe". In the case of FiO$_2$, pO$_2$ and pCO$_2$ a fourth score was added which represents the attribute "life-threatening". Body temperature used only three scores which represented the categories "normothermic", "mildly febrile" and "febrile".

The parameter FiO$_2$ was eventually removed from the analysis because of multicollinearity with high flow nasal cannula.

Table 13: Scoring for Clinical Events with Numerical Data Type

Score Name	Parameter	Score	greater than	less or equal
ssSpO2	SpO2	0	94%	100%
		1	90%	94%
		2	87%	90%
		3	0%	87%
ssHR	HR	0	-1th centile	90th centile
		1	90th centile	95th centile
		2	95th centile	99th centile
		3	99th centile	1000th centile
ssRR	RR	0	-1th centile	90th centile
		1	90th centile	95th centile
		2	95th centile	99th centile
		3	99th centile	1000th centile
ssBP_sys	BP_sys	0	-1th centile	90th centile
		1	90th centile	95th centile
		2	95th centile	99th centile
		3	99th centile	1000th centile
ssBP_dia	BP_dia	0	-1th centile	90th centile
		1	90th centile	95th centile
		2	95th centile	99th centile
		3	99th centile	1000th centile
ssTemp	Temp	0	36°C	37.8°C
		1	37.8°C	38.5°C
		2	38.5°C	41.5°C
ssABG_pH	ABG_pH	0	7.35	7.45
		1	7.25	7.35
		2	7.2	7.25
		3	6	7.2
ssABG_pO2	ABG_pO2	0	80 mmHg	1000 mmHg
		1	70 mmHg	80 mmHg
		2	60 mmHg	70 mmHg
		3	50 mmHg	60 mmHg
		4	0 mmHg	50 mmHg
ssCBG_pCO2	CBG_pCO2	0	30 mmHg	45 mmHg
		1	45 mmHg	55 mmHg
		2	55 mmHg	65 mmHg
		3	65 mmHg	75 mmHg
		4	75 mmHg	1000 mmHg
ssCBG_pH	CBG_pH	0	7.33	7.45
		1	7.25	7.33
		2	7.2	7.25
		3	6	7.2
ssCBG_pO2	CBG_pO2	0	60 mmHg	1000 mmHg
		1	50 mmHg	60 mmHg
		2	40 mmHg	50 mmHg
		3	30 mmHg	40 mmHg
		4	0 mmHg	30 mmHg

4.1.2 Categorical Event Type

Detailed analysis of all available clinical events revealed a group of text-based data fields describing the patient's current clinical state in a semi-structured way. These events were titled as follows: "Affect/Mood of Patient", "Level of Consciousness", "Skin Color", "Room Air", "Nasal Flaring", "Respiration Type", "Respiratory PEWS", "Respiratory Effort", "Respiratory Retraction", "Upper Airway Sounds", and "Breath Sounds".

For each data field the EHR software provided a list of pre-defined words or word combinations the user could choose from. The user was also permitted to enter free text. This unstructured part of the data entry was highlighted by the word "Other: ". The EHR software stored the text items delimited by a comma as a list in an alphanumerical data field (max. 255 characters).

Table 14 shows some examples of words and word combinations. The equal sign followed by a number is added for demonstration purposes. It illustrates which score was assigned to the text item. Text items without a number had a value of zero.

Table 14: List of Clinical Events Used for Severity Score with a Categorical Data Type

Name of Data Field	Sample of Text Items*
Affect/Mood of Patient	Awake, Crying, Interactive, Playful, Quiet, Sleeping, Uncooperative, Upbeat… Anxious=2, Fearful=2, Irritable=1, Other: restless=1
Level of Consciousness	Agitated, Alert, Arousable, Asleep, Quiet, Responsive… Irritable=1, Lethargic=2, Obtunded=1, Pt in respiratory distress=2, Restless=1
Skin Color	Flushed, Jaundiced, Normal for Ethnicity, Pale, Pink... Cyanotic=3, desat=3, Dusky=5, Mottled=1, Other: dusky when agitated=3
Room Air	Yes, No=3
Nasal Flaring	Absent, None Noted. Mild=1, Moderate=2, Present=1, Severe=3
Respiration Type	Regular. Shallow=1, subcostal retraction=1, Irregular=1, nasal flaring=2, Tachypnea=2, Dyspnea=2, increased work of breathing=2, Bradypnea=2, Apnea=3, Grunting=3, Agonal=4, Gasping=4, Other: head bobbing=3, With Ventilator=5
Respiratory PEWS	0. No retractions, 0. Within normal parameters for age. 1. Any assisted ventilation=2, 1. Any supplemental oxygen required to maintain normal =1, 1. Mild retractions=2, 1. O2 > 21% FiO2 required to maintain SpO2=1, 1. Using accessory muscles=3, 2. Moderate retractions=4, 3. Severe retractions and grunting=4
Respiratory Effort	No Distress, None, Normal, Comfortable.
Respiratory Effort - ED	Increased=1, Mild Distress=2, Moderate Distress=3, Retractions=1, Severe Distress=4
Respiratory Retraction	None. Mild=2, Moderate=4, Severe=6 Subcostal=1, Substernal=1, Intercostal=2, Supraclavicular=3, Suprasternal=3 All Muscles Used=4, head bobbing=3,
Upper Airway Sounds	Clear. wheezing=1, Grunting=2
Breath Sounds Bilateral Breath Sounds - Bilateral, Post Tx Breath Sounds - Bilateral, Pre Tx Breath Sounds - Left Breath Sounds - Right	Absent, Anterior, Before Treatment, Bowel Sounds, Clear, Coarse Crackles, Diminished in Bases, Fine Crackles, Focal, Moist Crackles, Pleural Rub, Posterior, remains diminished., Rhonchi, Scattered, Squeaks… Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Focal=1, Wheezing - Diffuse=2, Wheezing=2

* see Table 30 and Table 31 in the appendix for complete list of text items

4.2 Visualization of the Severity Score

Based on the steps mentioned above the dataset was scanned for available elements of the clinical severity score. In the next step, a data visualization tool, developed by the author, assisted with examining the course of the disease. The tool was also convenient for checking outliers and errors of data entry. Fig. 9 illustrates how this application summarized all relevant parameters on one single screen. The timeline is provided in minutes, i.e. 1440 minutes represented one day. The total length of stay for the patient depicted in Fig. 9 was 7,430 minutes which equated to approximately five days.
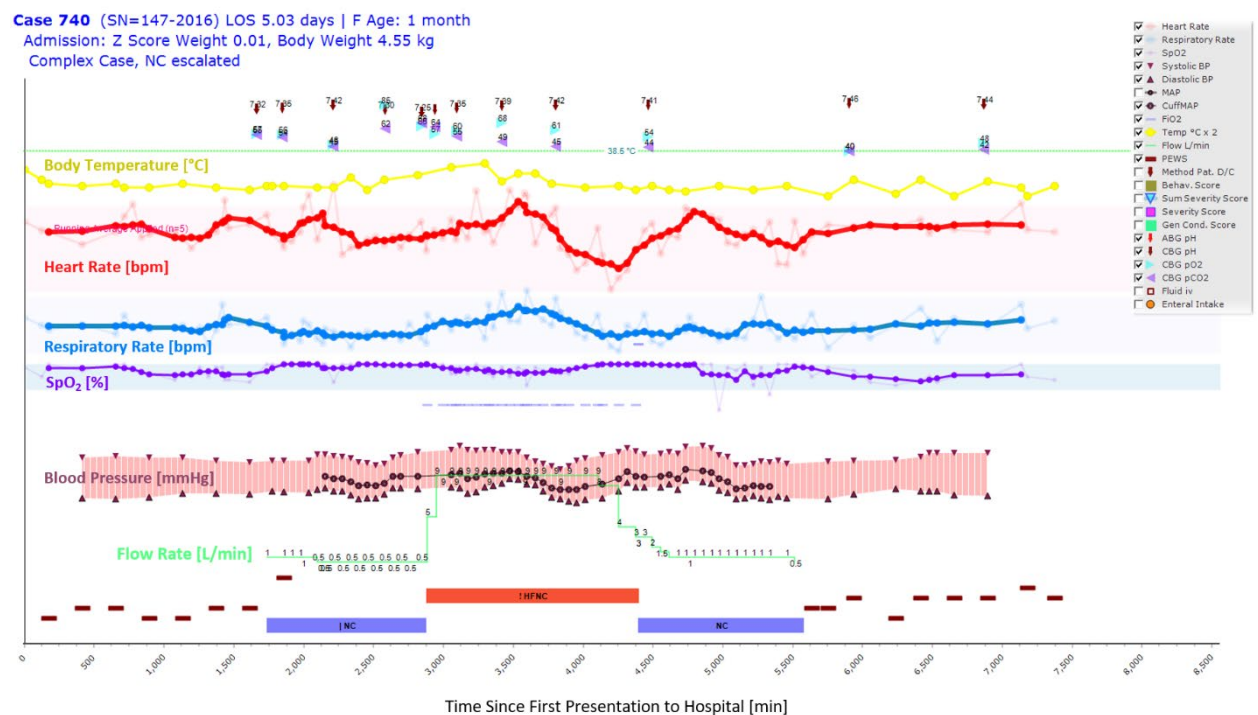
Fig. 9: Patient Chart Without Severity Score

To make trends better visible, a running average over five consecutive data points was applied to heart rate, respiratory rate and SpO2. The underlying curve was still visible. This example shows that heart rate, respiratory rate and SpO2 were within age-adjusted normal limits depicted by the lightly coloured boxes.[65,66] The green horizontal line at the top of the graph indicated the temperature limit for significant fever (>38.5 °C). The highest point on the yellow temperature curve represented a temperature of 37.9 °C. The green curve at the bottom of the page showed the nasal flow in litre per minute. The purple horizontal bar at the bottom represented the time-period of oxygen application via standard nasal cannula (NC). The flow was between 0.5 and 1 Litre per minute for that NC period. The following red bar on the Gantt chart represented high flow nasal cannula (HFNC) which was applied with up to nine Litre per minute, which was, in this case, in accordance with the recommended flow for HFNC of two Litre per minute and kilogram body weight.

The next chart demonstrates how the new data-driven severity score adds a new dimension to the interpretation of this patient's data. The sample frequency was set to follow that of the heart rate. This was too frequent. It caused the curve to sometimes undulate between extreme values like two and eighteen within just one hour. Reducing the sampling frequency would safely solve this problem because the score was cumulative, adding up every single individual score. To achieve more clarity on the trend, a running average was also applied to the respiratory score.
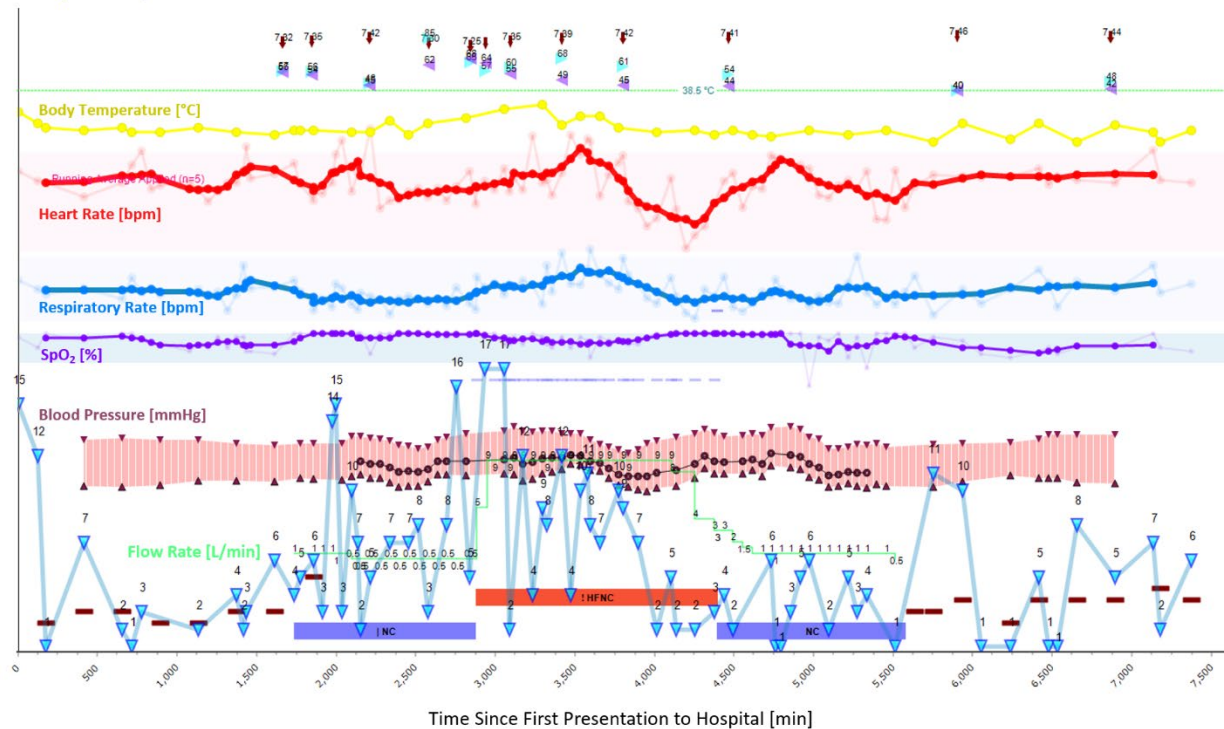
Fig. 10: Patient Chart with Raw Severity Score

The graph in Fig. 11 depicts how the severity score, incorporating numerical and categorical (=text) items, made interpretation of the course of illness much easier by drawing the undulating curve of the severity score.

Subsequently, visualisation of the severity score provided an interpretation as follows: The one-month old infant presented to hospital with an elevated score which subsequently decreased for one day. On the following day, a gradual increase of the severity score lead to the application of low-flow (0.5-1 L/min) oxygen via nasal cannula. Initially, the score continued to rise but then improved over the following six hours. Eventually, it increased to levels too high for standard care. The treating team changed from low flow to high flow therapy. Over the course of one day, the patient recovered, as reflected by the curve of the severity score. Eventually, after one more day on low-flow therapy and another day on spontaneous breathing, the patient was discharged home.
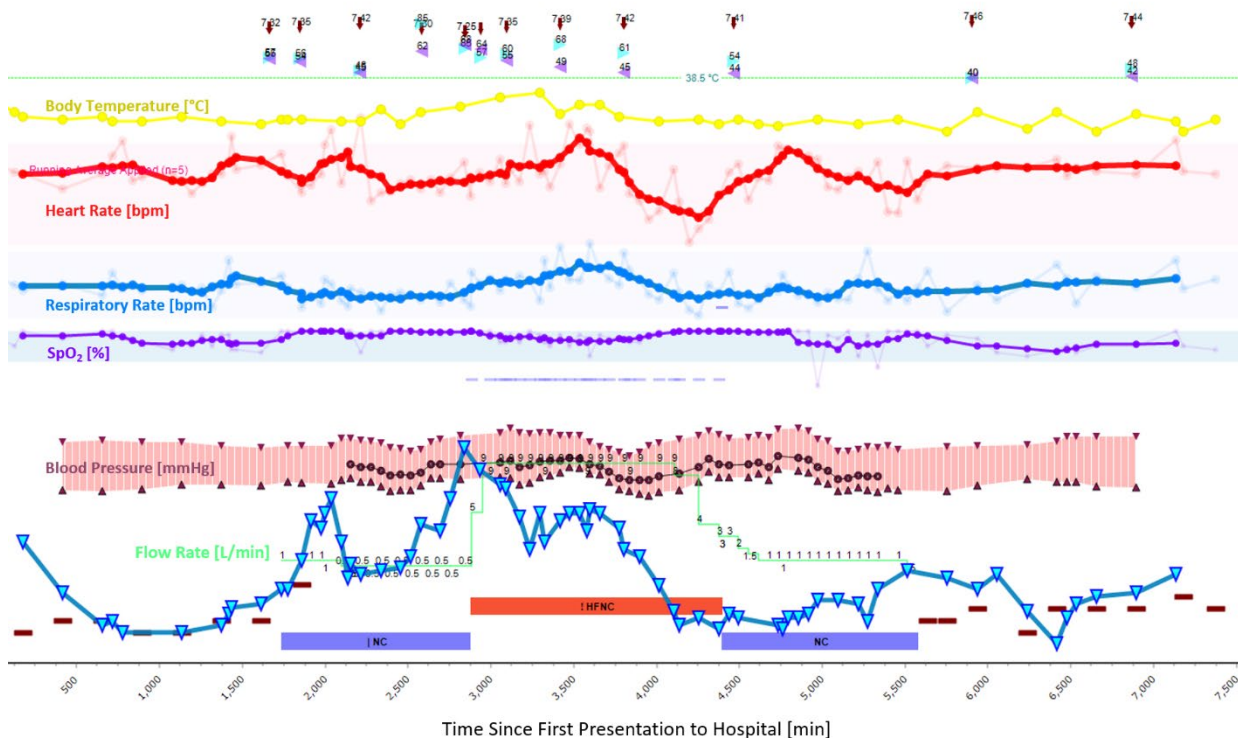
Fig. 11: Patient Chart with Running Average (5-point) of the Severity Score

A complete list of all 592 charts can be viewed online (Links: Simple Cases and Complex Cases, password: avb2018)

4.3 Assessment of Clinical Relevance

Based on the new score dataset, the levels of severity scores at different stages of the disease were examined. The aim was to correlate the score with outcome and high flow therapy. A computer algorithm identified the beginning of the recording period (hospital admission) and the beginning and end of the respiratory support period for each case. Six time periods were defined and analyzed as follows:

1 .... First six hours of the hospital stay
2 .... Second six hours of the hospital stay
3 .... 120 min before and 20 min after the start of non-invasive ventilation

The severity score was extracted from three equally distributed time windows within each episode:

4 .... 1st Third of non-invasive ventilation
5 .... 2nd Third of non-invasive ventilation
6 .... 3rd Third of non-invasive ventilation
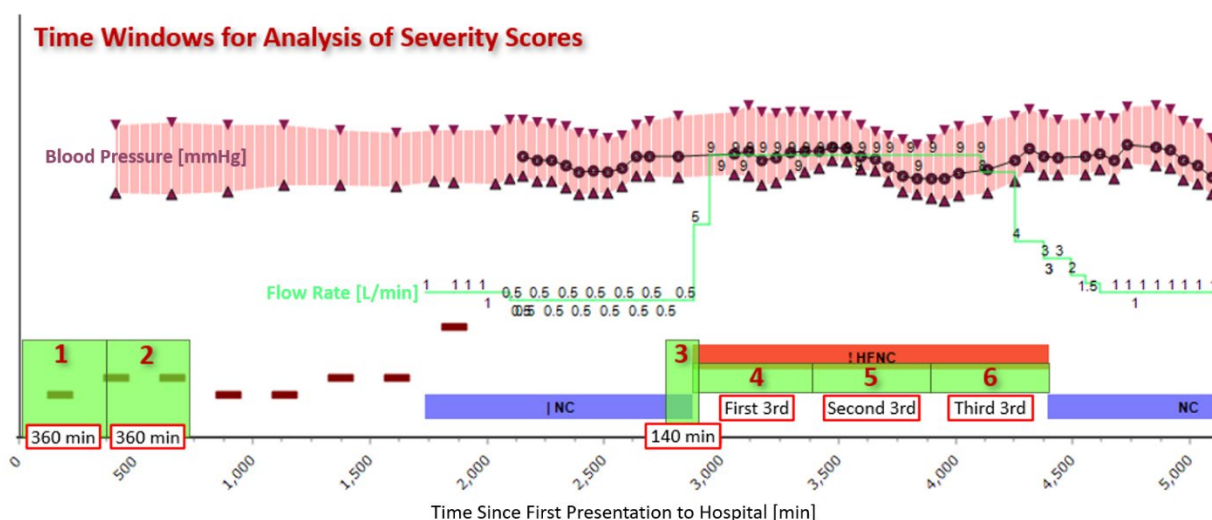
Fig. 12: Time Periods for Severity Score Analysis

Because the range of the score lied between zero and 300 a log transformation was applied which resulted in a score between 0 and 6. For each individual item the score was calculated and subsequently the average over the three-hour window was used. The final severity score was generated by calculating the sum of all individual scores. The results of the severity score assessment based on the described sampling windows can be found in the chapters five to eight.

4.4 Discussion

Detailed item analysis, performed by two experienced paediatric intensive care physicians, identified twenty-two items that were considered relevant for describing a respiratory severity score. The items were twelve routinely measured items and ten categorical items which were based on clinical observations performed by clinicians, i.e. doctors, nurses, and respiratory technicians. Preliminary visual analysis of patient charts supported the hypothesis that a mixed, data-driven and manually assessed, respiratory severity score was potentially helpful to assess the patients state of illness and better understand the course of illness.

Even though it was not the main objective of this study to validate a newly created data-driven respiratory severity score, it provided five out of fifteen quality criteria of the Bekhof Quality Criteria for Dyspnoea Scores (Appendix C).[34] It was beyond the scope of this study to examine every aspect of score validation. This study focussed on how retrospective data could be used to describe severity and guide future research. Out of the complete dataset available for each patient's hospital stay, six time periods were selected to answer the research questions of this study (Section 1.1.2, p. 16): first and second 6-hour period beginning from first recorded event (emergency department of hospital ward), 140-minute time window before commencement of non-invasive ventilation, and 1st, 2nd and 3rd third of non-invasive ventilation with variable length.

Only a few studies examined the ability of the respiratory severity score to help with therapeutic decisions and predict outcome such as length of stay or length of oxygen requirement. Golan-Tripto and colleagues examined 50 children with acute viral bronchiolitis. They evaluated the modified Tal Score (MTS). Inter- and intra-observer reliability was found to be good, however, the first score obtained upon admission was only a fair predictor of oxygen requirement at 48 h, and length of stay at 72 h.[69] Their findings highlighted the difficulty of distinguishing between severity and outcome. The fact that a measurement instrument describes severity accurately and reliably does not necessarily mean that it can predict outcome. In a routine clinical setting, other factors like age, z score of body weight, viral agents, and co-morbidities play an important role as predictors of outcome.

Recently, Rodriguez-Martinez and colleagues performed a systematic review on all available severity scores.[38] They found that 60 out of 77 studies used a respiratory severity score for evaluating the efficacy of specific therapeutic interventions. Inter- and intra-observer reliability was often not assessed. According to Rodriguez-Martinez et al., the score developed by Marlais and colleagues was the best available instrument. Its clinical usefulness was limited because it only evaluated risk of hospital admission.[39]

McCallum and colleagues used two scoring systems (Tal and modified Tal score) on 115 children with viral bronchiolitis to assess its predictive value for need of supplementary oxygen at

12h and 24h post assessment. The area under receiver operating curve (AUC ROC) was found to be too low (0.69 and 0.75) for predicting oxygen requirement.[70,71]

Caserta and colleagues used principal component analysis to develop a global respiratory severity score for RSV infections in infants. Nine clinical variables were identified and weighted to produce a composite global respiratory severity score. The calculations were made on data collected at three time points from hospitalized patients (n=84) and patients seen as outpatients (n=55). The study did not attempt to predict outcomes in individual patients or measure changes of severity during the course of illness. Its main focus was on quantification of overall illness severity which was used to predict hospitalization.[72]

The following chapters five to eight demonstrate how a data-driven respiratory severity score might be useful to predict hospital length of stay, assist with the treatment decision between low flow and high flow therapy, help to assess the treatment effect of high flow therapy, and finally, predict success and failure of non-invasive ventilation.

5 Prediction of Length of Hospital Stay

Length of stay represents the most relevant outcome variable for infants with acute viral bronchiolitis. It is of great interest for patients and their accompanying parents, and for the hospital to assist with planning of staff and resources. For this reason, this study used length of stay as the primary outcome variable.

In this chapter, machine learning methods are used to identify the most relevant variables that predict length of stay (feature selection). Correlation studies and calculation of odds ratios helped to determine the relationship between independent variables and hospital length of stay.

5.1 Total length of Stay (tLOS)

One difficulty of using length of stay as the primary outcome variable is determining true outcome. As discussed earlier (Section "*Length of Stay",* p. 33), the total length of hospital stay was calculated in minutes from the first recorded event until discharge. Brooks and colleagues discussed the effect of discharge criteria on research results.[73] Most commonly, readiness for discharge is defined by the following variables: normal SpO_2 in room air, no respiratory distress, and normal food intake for 6 or 12 hours. Unfortunately, readiness for discharge could not be determined due to insufficient data.

5.2 Feature Selection

Feature selection, also known as **variable selection**, **attribute selection** or **variable subset selection**, is the process of selecting a subset of relevant features (synonyms: items, variables, predictors) for use in model construction. It helps to determine the features that have the most important influence on the response variable. Feature selection techniques simplify models to make them easier to interpret. They reduce training times, avoid the problems of high dimensionality, and enhance generalization by reducing overfitting.[74,75] Reducing the dimensionality of the data can often be done without significant loss in performance and predicting power.

Table 15: R Packages Used in this Study

Model	Full Name	R Package
Glm	Generalized Linear Models	stats (R core) [76]
LASSO (MSE)	Least Absolute Shrinkage and Selection Operator	glmnet[77]
BORUTA	Wrapper approach built around a random forest	boruta[78]
BART	Bayesian Additive Regression Trees	bartMachine[79,80]
XGBOOST	Extreme Gradient Boosting	xgboost[81]

5.2.1 Applied Feature Selection with R

This chapter describes in a first step which factor and clinical events influenced length of stay most. In total, 28 variables were examined by five different methods: glm, LASSO, Boruta, BART, XGboost. Due to the retrospective nature of this study's observational dataset, the results

had to be interpreted with caution because some variables contained duplicate content, some occurred only in small numbers, and some were distributed unevenly. In a second step, the influence of the calculated severity score on the length of stay was estimated (*sumSS*, see previous chapter: *Severity Score for Respiratory Distress*, p.42). This was done in conjunction with eight other variables.

The following code demonstrates the key statements written in R. The complete R code can be found online: http://ckcdata.com/R-code/Feature-Selection-tLOS-v1.0.html, http://ckcdata.com/R-code/Feature-Selection-tLOS-sumSS.html.

Table 16: R Functions Used for Feature Selection

Method	Key statement
glm	glm(tLOS ~ ., data = myData, family = gaussian, maxit = 100)
LASSO	cv.glmnet(x, y, alpha = 1, family = "gaussian", type.measure = "mse", nfolds = 10)
BORUTA	Boruta(HFNC ~ ., data = myData, doTrace = 1)
BART1	investigate_var_importance
BART2	var_selection_by_permute(BART_fit_PS, num_reps_for_avg = 10, num_permute_samples = 100, num_trees_for_permute = 20, alpha = 0.05, plot=TRUE)
XGBOOST	

5.2.2 Glm (Generalized Linear Model)

The R function *glm* (R package stats) is used to fit generalized linear models, specified by giving a symbolic description of the linear predictor and a description of the error distribution.[76]

In relation to total length of stay, the following excerpts of the R code (Code 1) show that applying a Gaussian model (assuming normal distribution) returns significant p values for the z-score of the body weight (zWt), the age at admission in days (AGEREG_DAYS), RSV, and Metapneumo Virus. The negative estimates for zWt and AGEREG_DAYS indicate a negative correlation, e.g. for every increase of AGEREG_DAYS by one day tLOS decreases by 3.841 minutes, for every increase of zWt by one, tLOS decreases by 382.207 minutes. According to this model, detection of RSV increases tLOS by 484.902 min, and detection of Metapneumo Virus increases tLOS by 1555.471 min, which is more than one day.

```
## myData: 607 obs. of 29 variables

mFormula
tLOS ~ zWt + AGEREG_DAYS + RSV + virusRhinoEntero + virusMetapneumo +
  virusAdeno + ssMood + ssConscious + virusParainf3 + ssABG_pH +
  ssSkinColor + ssRoomAir + ssSpO2 + ssHR + ssRR + ssBP_sys +
  ssBP_dia + ssTemp + ssCBG_pCO2 + ssCBG_pH + ssCBG_pO2 + ssNasalFlaring +
  ssRespType + ssRespPEWS + ssRespEffort + ssRespRetraction +
  ssUpperAirwaySounds + ssBreathSounds

glm_model <- glm(mFormula, data = myData,  family=gaussian, maxit = 100)
summary(glm_model)

Deviance Residuals:
    Min        1Q    Median        3Q       Max
 -4305.0   -1830.1    -556.9    1347.1    8319.7
```

```
Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      6367.419    245.045  25.985  < 2e-16 ***
zWt              -382.207     71.958  -5.312 1.55e-07 ***
AGEREG_DAYS        -3.841      1.089  -3.527 0.000453 ***
RSV               484.902    227.081   2.135 0.033151 *
virusMetapneumo  1555.471    428.257   3.632 0.000306 ***
```
...

Code 1: Summary of GLM model, only showing significant variables.

5.2.3 LASSO (Least Absolute Shrinkage and Selection Operator)

The R package glmnet provides LASSO regularisation which selects the most significant variables and eliminates less important ones. The function cv.glmnet performs a grid search to find the optimal value of Lambda.
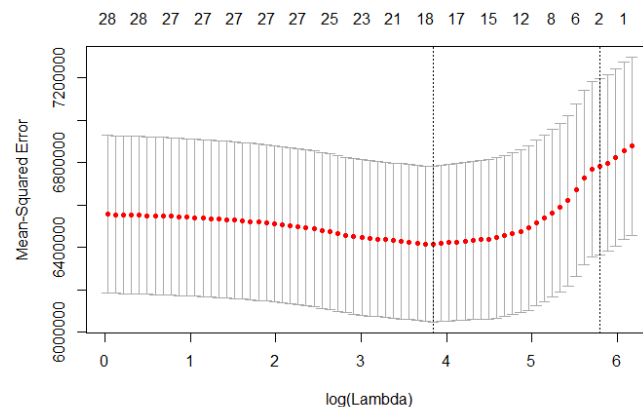

```
avb.lasso.cv <- cv.glmnet(x, y, alpha = 1,
                      family = "gaussian", type.measure = "mse", nfolds = 100)
plot(avb.lasso.cv)
(lambda_min <- avb.lasso.cv$lambda.min)
(lambda_1se <- avb.lasso.cv$lambda.1se)
rc <- coef(avb.lasso.cv, s=lambda_1se)
summ <- summary(rc)
ImportantVarMSE <- data.frame(VarName = rownames(rc)[summ$i], Coeff = summ$x)
print(ImportantVarMSE)


VarName          Coeff
(Intercept)      6501.833796
zWt              -99.056945
ssRespEffort     5.386503
```

Code 2: LASSO regularisation (mean squared error)

The result of LASSO lists the z score of the body weight (zWt) and respiratory effort (ssRespEffort)[1] as the most important independent variables.

5.2.4 BORUTA

Boruta[2] is a random forest-based method that iteratively removes variables that are statistically less relevant than random probes which are introduced as artificial noise variables.[74,78]

```
library(Boruta)
library(mlbench)

set.seed(1)
feature.selection <- Boruta(tLOS ~ ., data = md, doTrace = 1)
table(feature.selection$finalDecision)
```

---

[1] ssRespEffort - keywords: "Increased"=1, "Mild Distress"=2, "Moderate Distress"=3, "Other: Grunting"=4, "Other: Head bobbing"=3, "Retractions"=1, "Severe Distress"=4, "tachypneic"=1
[2] Boruta is a god of the forest in the Slavic mythology

```
getConfirmedFormula(feature.selection)
par(mar=c(9,5,4,1)+.1)
plot(feature.selection,colCode=c('green','yellow','red','blue'),sort=TRUE,
    whichRand=c(TRUE,TRUE,TRUE),col=NULL,las=2,
    #xlab='Attributes',
    xlab='',
    ylab='Importance')
```

Code 3: BORUTA algorithm
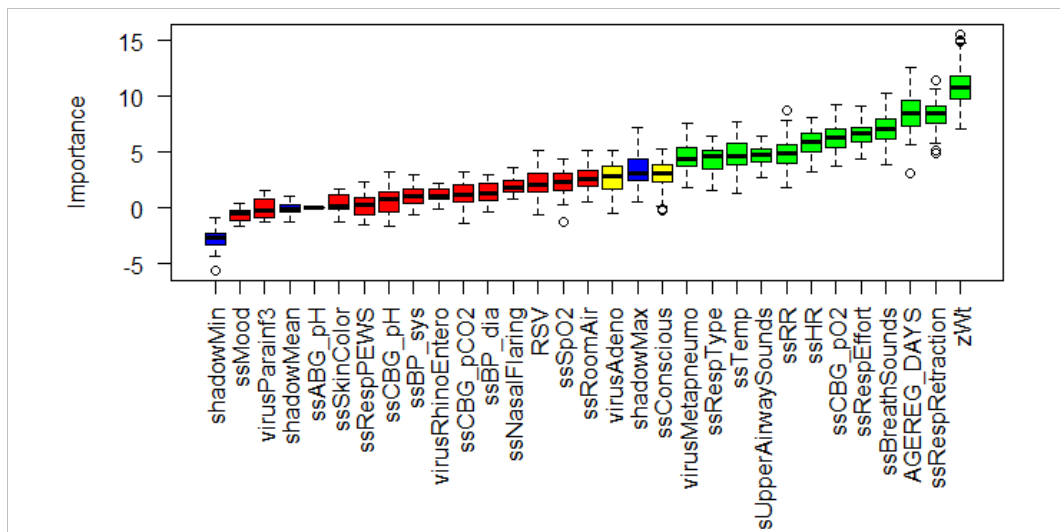
Fig. 13. BORUTA Feature Selection

The result of the BORUTA algorithm lists the z score of the body weight (zWt), respiratory retraction and admission age as the three most important variables.

5.2.5   BART (Bayesian Additive Regression Trees)

The R package *bartMachine* provides an advanced implementation of Bayesian Additive Regression Trees with expanded features for data analysis and visualisation.[80,82-85] This study used the function *bartMachine* to build a BART regression model.

```
library(rJava)
library(bartMachine)
library(bartMachineJARs)
library(rms)
library(cvTools)

X <- md[,1:c0]
Z <- md$tLOS

set_bart_machine_num_cores(4)
bart = bartMachine(X=data.frame(X), y=Z, num_trees=200)
summary(bart)

bartMachine v1.2.3 for regression

training data n = 607 and p = 28
built in 6.4 secs on 4 cores, 200 trees, 250 burn-in and 1000 post. samples

sigsq est for y beforehand: 5925622.212
avg sigsq estimate after burn-in: 5896884.42731

in-sample statistics:
 L1 = 1158259.51
 L2 = 3340431889.86
 rmse = 2345.89
 Pseudo-Rsq = 0.1961
p-val for shapiro-wilk test of normality of residuals: 0
p-val for zero-mean noise: 0.96691


investigate_var_importance(bart, type = "splits",
                           plot = TRUE, num_replicates_for_avg = 5, num_trees_bottleneck = 20,
```

```
                            num_var_plot = Inf, bottom_margin = 10)
```
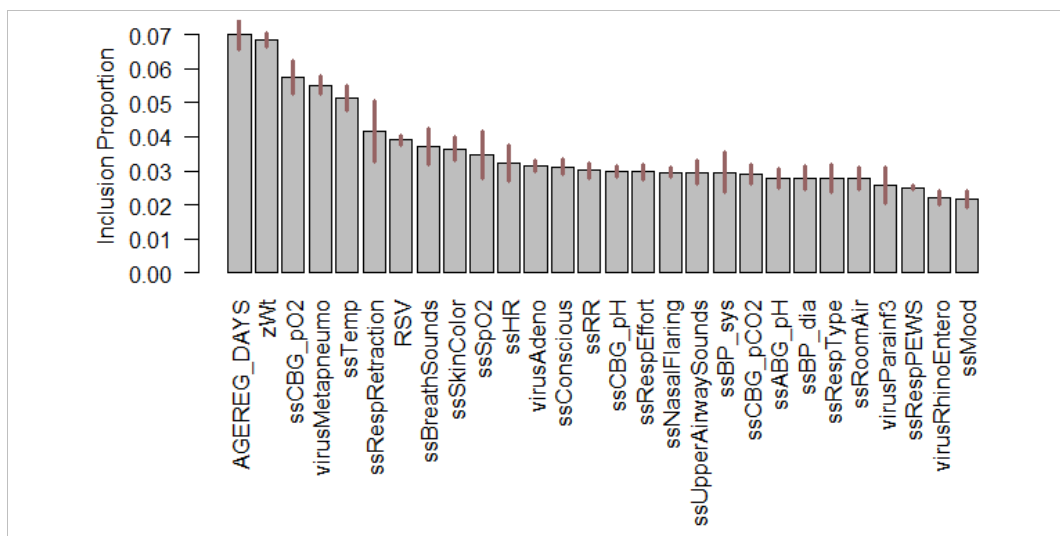

Fig. 14. BART Feature Selection

```
vs_bart <- var_selection_by_permute(bart, num_reps_for_avg = 10, num_permute_samples = 100,
                         num_trees_for_permute = 20, alpha = 0.05, plot=FALSE)
vchoose_bart <- vs_bart$important_vars_local_col_nums
vchoose_bart

print(names(X[,vchoose_bart]))

[1] "AGEREG_DAYS"    "zWt"            "ssCBG_pO2"      "virusMetapneumo" "ssTemp"
```

The function *var_selection_by_permute* uses three thresholding methods introduced by Bleich et al.[85] The result contains the admission age, the z score of the body weight (zWt), and pO_2 measured via capillary blodd gas (ssCBG_pO2) as the three most important variables.

5.2.6   XGBOOST (Extreme Gradient Boosting)

XGBoost is an implementation of gradient boosting machines initially created by Tianqi Chen.[86] It supports several forms of gradient boosting. In this paragraph, XGBOOST is used to determine the most important variable by 10 fold cross-validation

```
param <- list(  objective         = "reg:linear",
                booster           = "gbtree",
                eval_metric       = "error",
                eta               = 0.1,
                max_depth         = 3,
                subsample         = 0.5,
                colsample_bytree  = 1,
                min_child_weight  = 1,
                gamma             = 0.5
)
param

x <- as.matrix(md[, 1:c0])
y <- as.matrix(md[, c1])
#y <- ifelse(md$HFNC == "1", 1, 0)
md.mat <- xgb.DMatrix(data=x, label = y)


set.seed(1)
md.xgb.fit <- xgb.train(params = param, data = md.mat, nrounds = 75)
md.xgb.fit

pred <- predict(md.xgb.fit, x)
summary(pred)
```

```
# head(pred)
# head(y)
impMatrix <- xgb.importance(feature_names = dimnames(x)[[2]], model = md.xgb.fit)
impMatrix

xgb.plot.importance(impMatrix, main = "Gain by Feature")
```
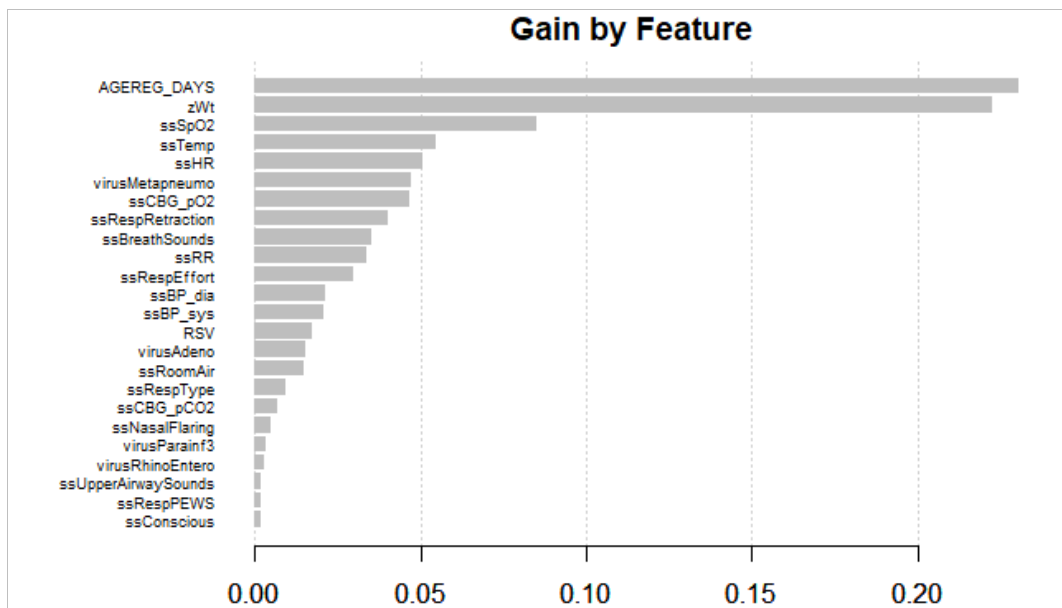

Fig. 15: XGBOOST – Gain by Feature for Total Length of Stay

The two most important features predicting total length of stay, selected by the XGBOOST package, were admission age and z score of the body weight (zWt).

5.2.7   Summary of Feature Selection

Table 17 summarises the results of five different methods to determine the most important features for predicting length of stay. The table lists only the first eight selected features. Most algorithms recognised admission age and z score of the body weight as the most important features.

Table 17: Results of Feature Selection for Length of Stay – ss_all

	VarName	glm	LASSO MSE	BORUTA rank	BART Rank	XGBOOST	Count Selected
1	zWt	***	1	1	2	2	5
2	AGEREG_DAYS	***		3	1	1	4
3	virusMetapneumo	***			4	6	3
4	ssCBG_pO2			6	3	7	3
5	ssRespRetraction			2	6	8	3
6	ssBreathSounds			4	8		3
7	RSV	*			7		2
8	ssHR			7		5	2
9	ssRespEffort		2	5			2
10	ssTemp				5	4	2
11	ssRR			8			1
12	ssSpO$_2$					3	1

As discussed in chapter 4 (*Severity Score for Respiratory Distress*, page 42), manual inspection of the cases revealed that many items of the described severity score were not equally recorded, e.g. in simple cases many clinical parameters and observations were missing. In addition, the content of some of the individual scores (prefix "ss") was overlapping. Therefore, the sum of

the individual scores (sumSS) was used. The full script can be found online: 5.2-Feature-Selection-tLOS-sumSS-v1.1.html.

To assess the influence of the respiratory severity score at the time of the treatment decision for high flow nasal cannula or standard therapy, the severity score of time period 3 ("sumSS"; 120 min before and 20 min after the start of non-invasive ventilation; chapter 0, page 50) was used to perform a second run of the five aforementioned feature selection tools. The results are shown in Table 18. Fig. 16 depicts a graphical representation for extreme gradient boosting (booster = gbtree).

Table 18: Results of Feature Selection for Length of Stay – sumSS and five other variables

	VarName	glm	LASSO MSE	BORUTA rank	BART1 rank	XGBOOST	Count Selected
1	sumSS	***	+	5	4	3	5
2	zWt	***	+	1	2	2	5
3	AGEREG_DAYS	***	+	2	1	1	5
4	virusMetapneumo	***	+	3	3	4	5
5	RSV	*	+			5	3
6	virusAdeno			4	5	6	3
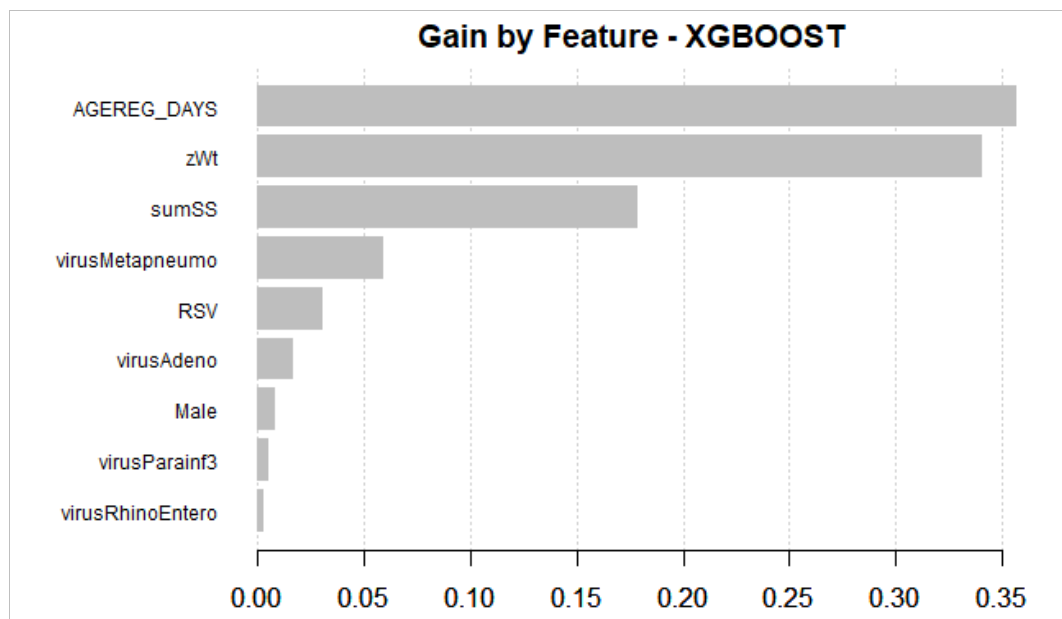
Fig. 16: Feature Selection Length of Stay; sumSS; Extreme Gradient Boosting (XGBOOST)

5.3 Correlation between z Score of Body Weight or Admission Age and Length of Stay

To test the correlation between z score of body weight or admission age and length of stay, linear regression and other models were used.
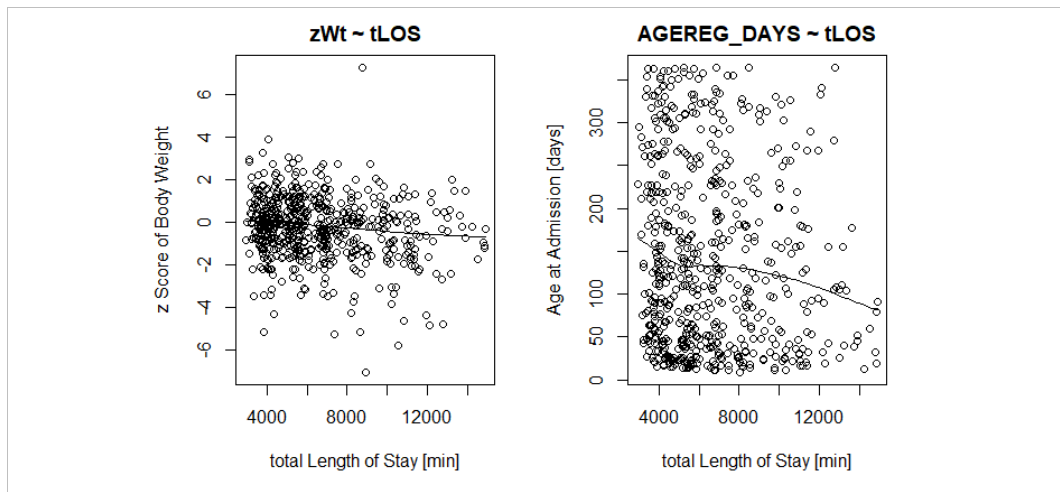
Fig. 17: Scatterplot of z Score Body Weight (zWT) and Age versus tLOS

The regression lines on the scatterplots in Fig. 17 suggest a negative correlation between the two variables. The left panel plots the relationship between the z score of the body weight and the total length of stay (tLOS) in minutes. The right panel shows the admission age in days in relation to total length of stay.

The following R code (Code 4) uses the *lm* function (linear model) to support the finding that the negative correlations between zWT and tLOS and AGEREG_DAYS and tLOS were statistically significant (p<0.001).

```
md <- md_ssNIV
linearModel <- lm(tLOS ~ zWt + AGEREG_DAYS, data=md)
summary(linearModel)

Residuals:
    Min      1Q  Median      3Q     Max
-4326.6 -2005.5  -648.6  1405.5  8191.9

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 6916.824    177.435  38.982  < 2e-16 ***
zWt         -367.025     73.124  -5.019 6.88e-07 ***
AGEREG_DAYS   -3.641      1.031  -3.531 0.000446 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2588 on 589 degrees of freedom (=Standard Error of the Estimate)
Multiple R-squared (=R2):  0.0513,    Adjusted R-squared:  0.04808
F-statistic: 15.93 on 2 and 589 DF,  p-value: 1.837e-07

confint(linearModel)
                2.5 %      97.5 %
(Intercept) 6568.342752 7265.306007
zWt         -510.641199 -223.409632
AGEREG_DAYS   -5.665821   -1.615943
```

Code 4: Result of linear regression for zWt and Age over Length of Stay

Using the R package *caret*, ten-fold cross-validation, repeated ten times, was applied to assess the performance of six different models (KNN, SVM, GLMNET, StepAIC, LM, GLM). Preliminary tests demonstrated that the Yeo-Johnson power transformation produced the best results as depicted in Fig. 18.[87,88]
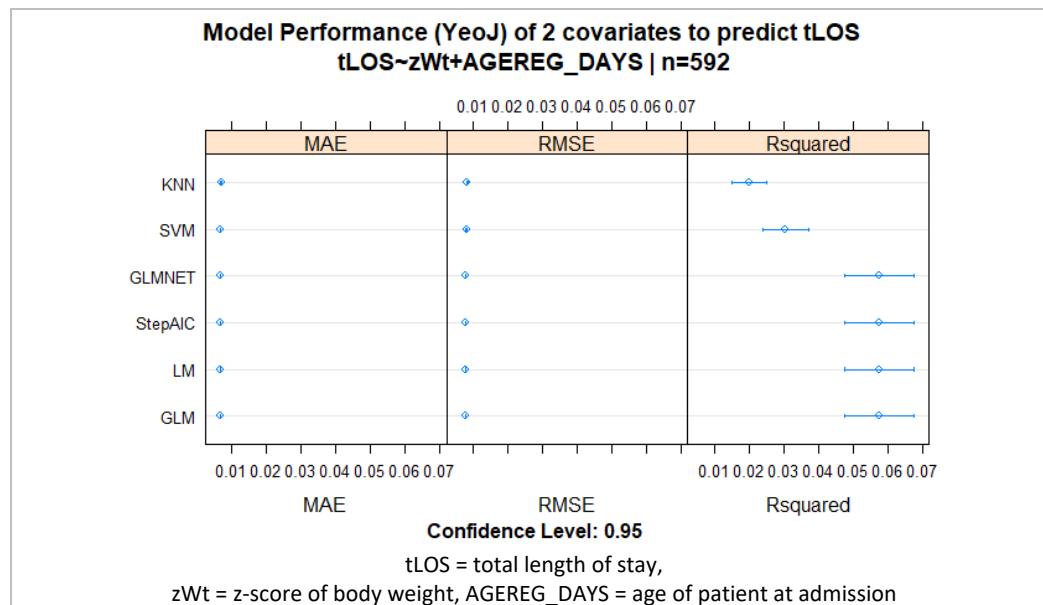
Fig. 18: Model Performance with z score body weight and age as predictors to predict total length of stay (Yeo-Johnson transformed values)

GLMNET, StepAIC, LM, GLM produced the highest R^2-values at approximately 5.77% which is still low suggesting high variability with low predictive precision.

5.4 Correlation between Severity Score and Length Stay

The severity score measured at the **time of the treatment decision** to apply either standard nasal oxygen or high flow nasal oxygen was used to create the following scatterplot (Fig. 19). It suggests that the severity score has a positive correlation with the total length of stay. This time, the finding was supported by fitting a linear regression model in R.
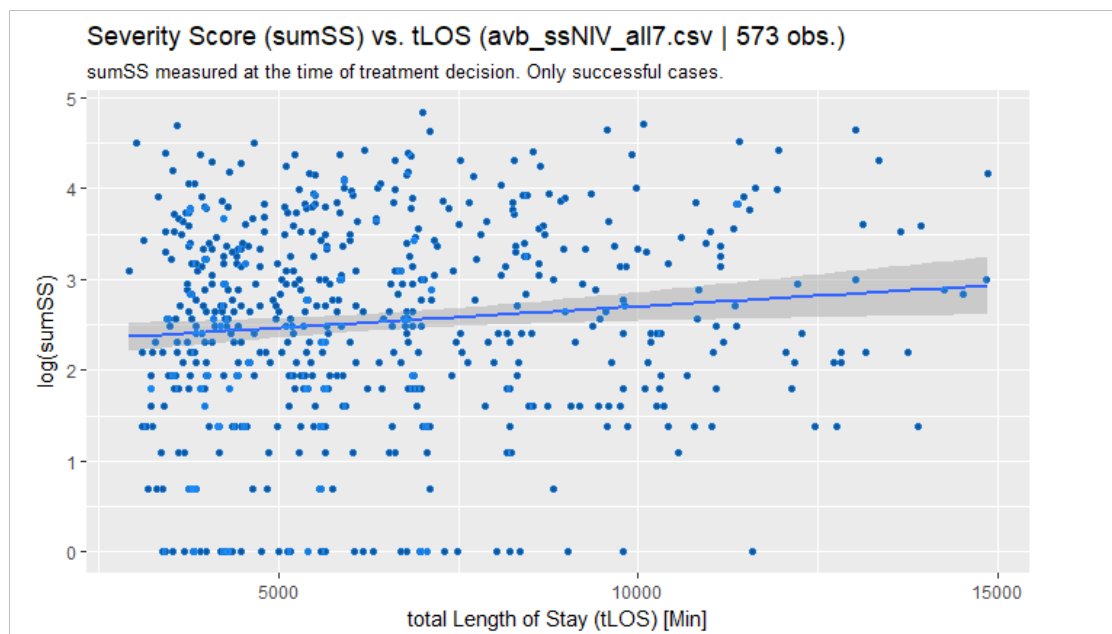
Fig. 19: Scatterplot Severity Score (sumSS) versus total Length of Stay

5.4.1 Test Model Performance – Unadjusted Predictors

The R package *caret* was used to test the performance of seven models by applying ten-fold cross validation repeated ten times (Table 19).

Table 19: Models used with *train* function (*caret*)

Short	Name	method value	Libraries
LM	Linear Regression	lm	
GLM	Generalized Linear Model	glm	
GLMNET	Lasso and Elastic-Net Regularized Generalized Linear Models	glmnet	glmnet, Matrix
SVM	Support Vector Machines with Radial Basis Function Kernel	svmRadial	kernlab
CART	Classification And Regression Trees	rpart	rpart
KNN	k-Nearest Neighbors	knn	
XGB	eXtreme Gradient Boosting	xgbLinear	Xgboost

The first run of the R code (see APPENDIX, page 130) used unadjusted data (*Transform=0*). The performance indicators R^2, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) demonstrated poor performance (Fig. 20).
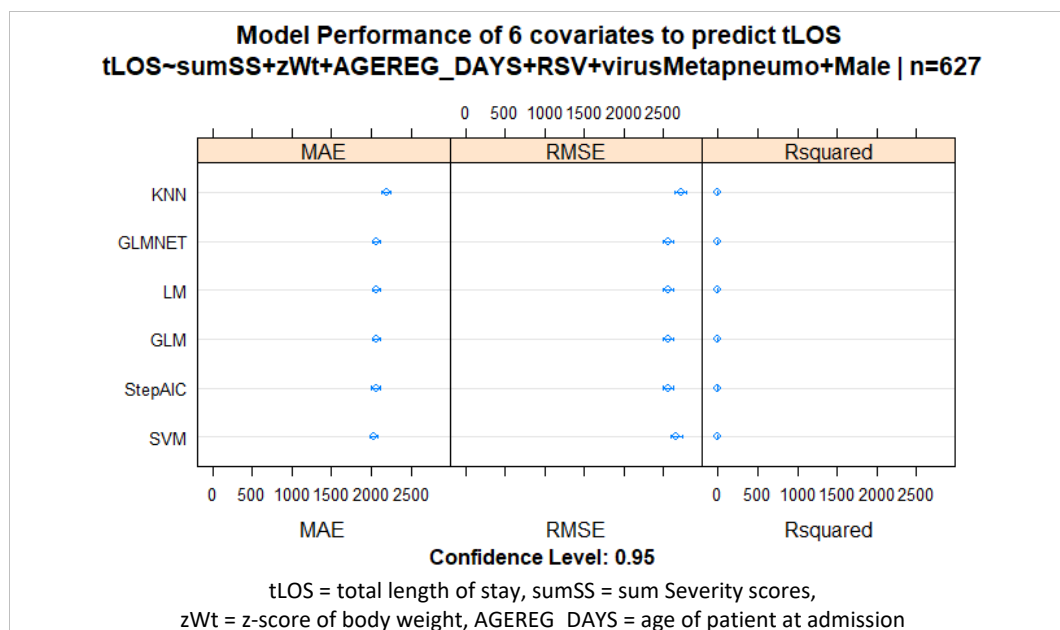
Fig. 20: Overview Model Performance for six covariates to predict total Length of Stay (original values)

Subsequently, the R package *gvlma* (Global Validation of Linear Models Assumptions) was applied to test whether the linear model with the unadjusted dataset conformed with the assumptions of linear regression.

```
library(gvlma)
mFormula <- tLOS ~ sumSS + zWt + AGEREG_DAYS
linearModel <- lm(mFormula, data=md)
summary(linearModel)
gvlma::gvlma(linearModel)

Call:
 gvlma::gvlma(x = linearModel)

                    Value p-value                  Decision
Global Stat       88.2064 0.00000 Assumptions NOT satisfied!
Skewness          77.0644 0.00000 Assumptions NOT satisfied!
Kurtosis           0.2913 0.58939    Assumptions acceptable.
Link Function      4.6849 0.03043 Assumptions NOT satisfied!
Heteroscedasticity 6.1659 0.01302 Assumptions NOT satisfied!
```

Code 5: Test Assumptions of Linear Regression for unadjusted values
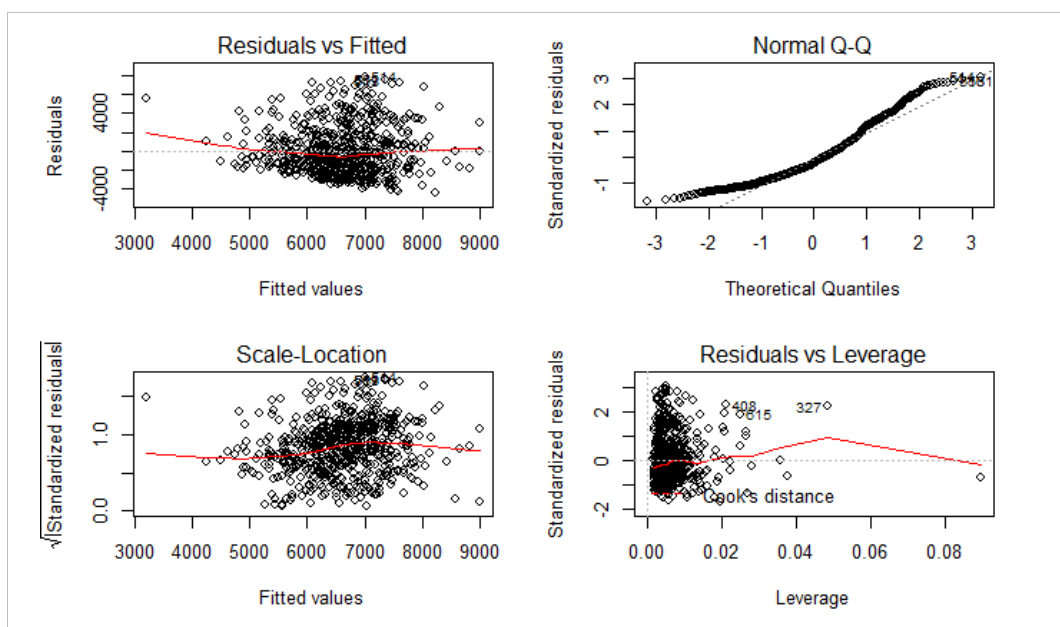
Fig. 21: Diagnostic Plots for Unadjusted Linear Model (tLOS ~ sumSS + zWt + AGEREG_DAYS)

The above code (Code 5) and Fig. 21 showed that almost all assumptions for linear regression were not met. Visual comparison of the distributions of the unadjusted values, log-transformed values and values transformed by the Yeo-Johnson Transformation helped to decide which transformation would help to create a nearly normal distribution (Fig. 22, Fig. 23). The Yeo-Johnson transformation was applied to the z score of the body weight (zWT), the age (AGEREG_DAYS), and total length of stay (tLOS). The log transformation was applied to the severity score (sumSS).
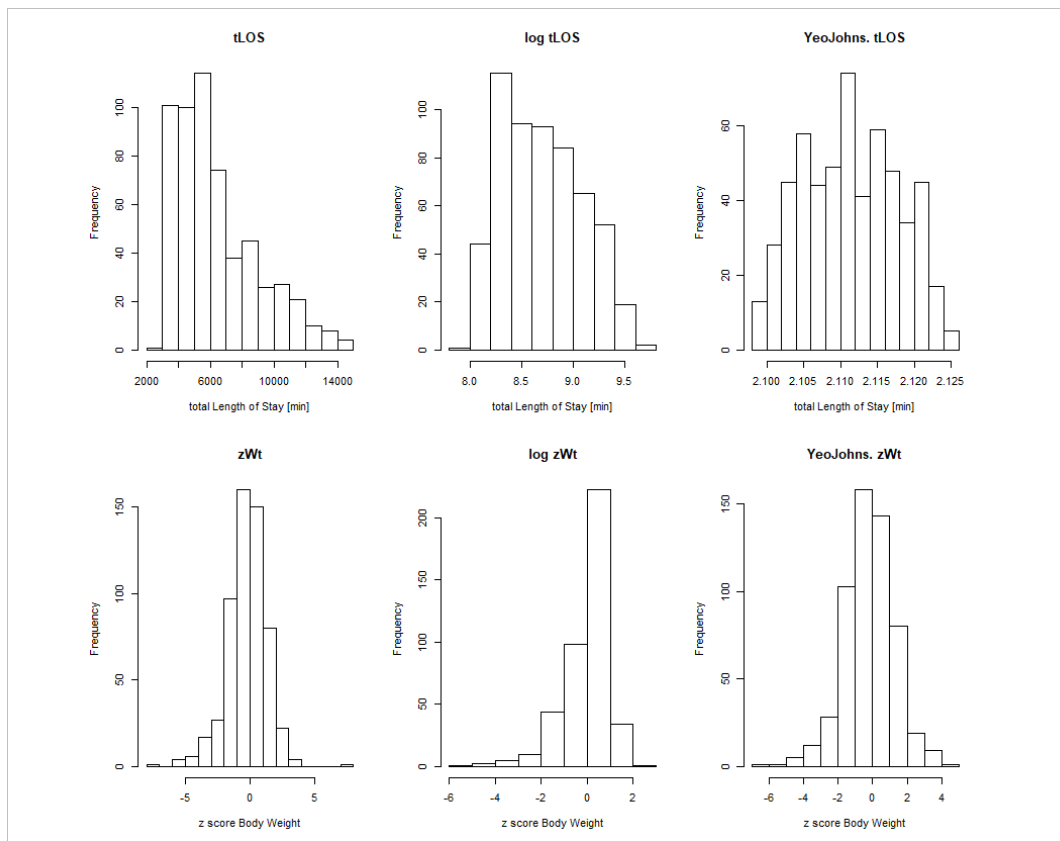
Fig. 22: Comparisons of three distributions (un-adjusted, log-transformed, YeoJohnson-transformed) for total length of stay (tLOS) and z score of the body weight (zWT)
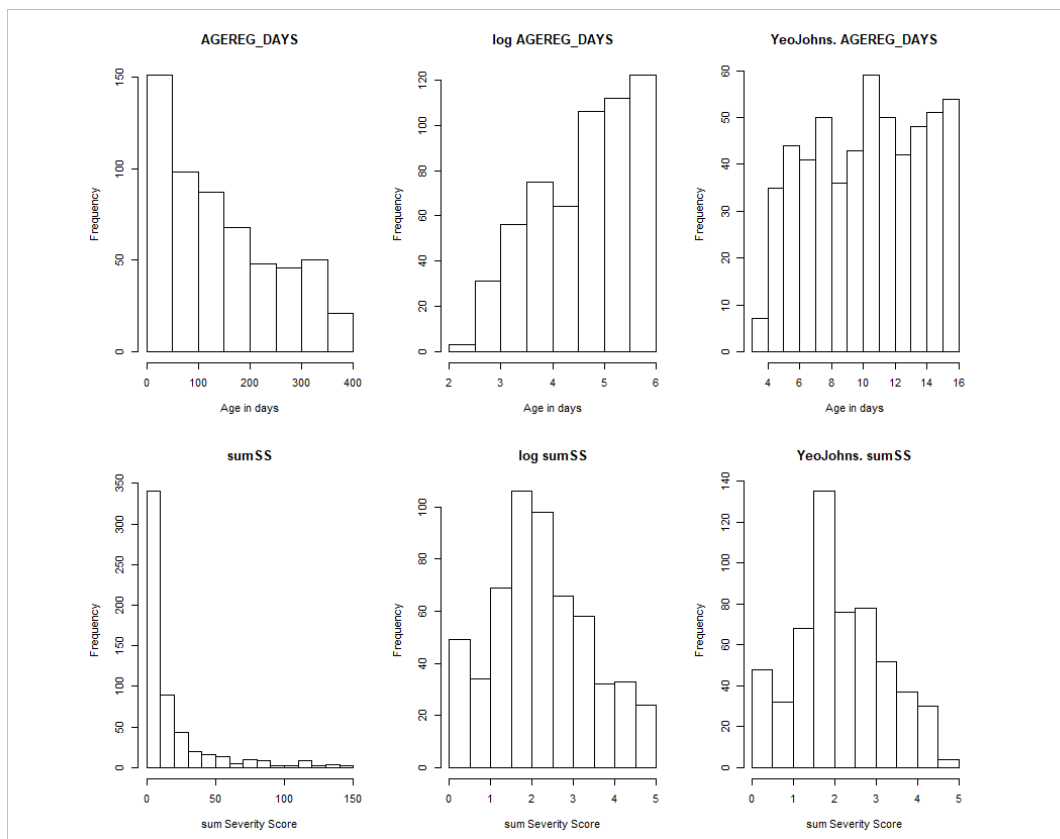
Fig. 23: Comparisons of three distributions (un-adjusted, log-transformed, YeoJohnson-transformed) for age at admission (AGEREG_DAYS) and respiratory severity score (sumSS)

After applying the transformation of the variables, the R function gvlma produced an improved result (Code 6, Fig. 24).

```
Call:
 gvlma::gvlma(x = linearModel)

                         Value    p-value                   Decision
Global Stat          26.11849  2.995e-05  Assumptions NOT satisfied!
Skewness              0.05911  8.079e-01      Assumptions acceptable.
Kurtosis             22.13278  2.544e-06  Assumptions NOT satisfied!
Link Function         3.75398  5.268e-02      Assumptions acceptable.
Heteroscedasticity    0.17262  6.778e-01      Assumptions acceptable
```

Code 6: Test Assumptions of Linear Regression for adjusted values (log and Yeo-Johnson)
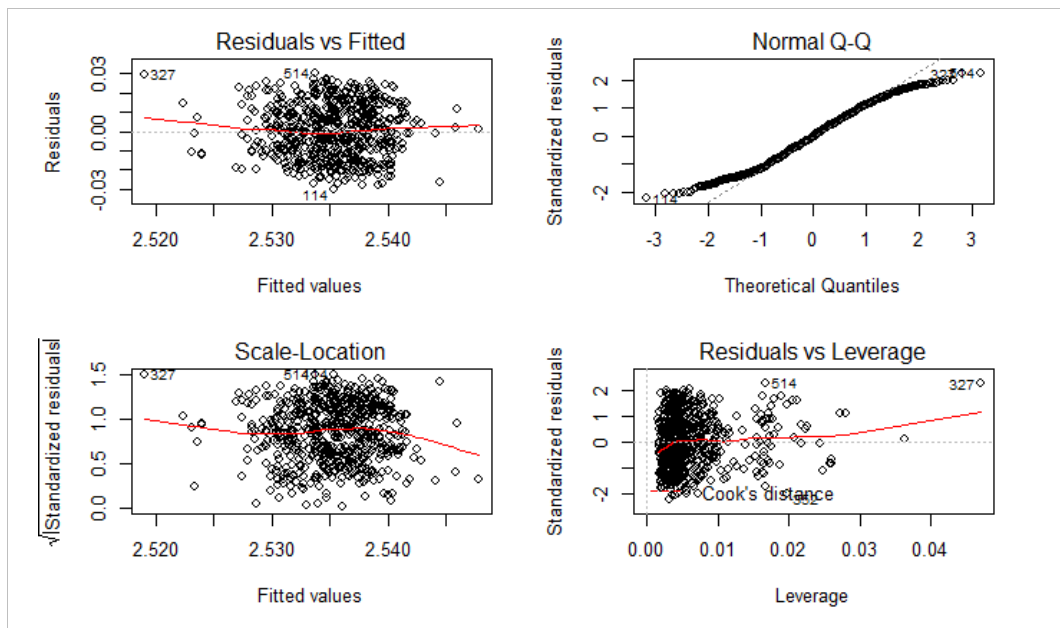
Fig. 24: Diagnostic Plots for Linear Model using transformed values (tLOS ~ sumSS + zWt + AGEREG_DAYS)

5.4.2 Test Model Performance – Adjusted Predictors

The second run of the R code (see APPENDIX, page 130) used transformed values (*Transform=1*) as described above. This time the performance indicators (RMSE, MAE, R^2) demonstrated improved performance (Fig. 25). LM, GLM, GLM with stepwise feature selection and GLMNET showed the highest R^2. RMSE and MAE were now positioned at the lower end. The summary command for the Generalized Linear Model with Stepwise Feature Selection (*method="glmStepAIC"*) lists highly significant p-values for all six covariates. The low adjusted R^2 (0.099) suggests high variability in the data with low predictive precision.

```
Generalized Linear Model with Stepwise Feature Selection

627 samples
  6 predictor

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 10 times)
Summary of sample sizes: 564, 563, 564, 563, 566, 564, ...
Resampling results:

  RMSE       Rsquared   MAE
  0.0133332  0.10068    0.01120534


Call:
NULL
```

```
Deviance Residuals:
      Min         1Q      Median         3Q         Max
-0.0293202  -0.0107851  -0.0001439   0.0099017   0.0314940

Coefficients:
                 Estimate Std. Error  t value Pr(>|t|)
(Intercept)     2.5348097  0.0021956 1154.493  < 2e-16 ***
sumSS           0.0017015  0.0004777    3.561 0.000397 ***
zWt            -0.0019080  0.0003723   -5.125 3.97e-07 ***
AGEREG_DAYS    -0.0007365  0.0001955   -3.768 0.000180 ***
RSV             0.0037310  0.0011435    3.263 0.001164 **
virusMetapneumo 0.0074635  0.0021586    3.458 0.000582 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.0001767557)

    Null deviance: 0.12153  on 626  degrees of freedom
Residual deviance: 0.10977  on 621  degrees of freedom
AIC: -3630.4

Number of Fisher Scoring iterations: 2
```

Code 7: Summary of Generalized Linear Model with Stepwise Feature Selection (*method="glmStepAIC"*)
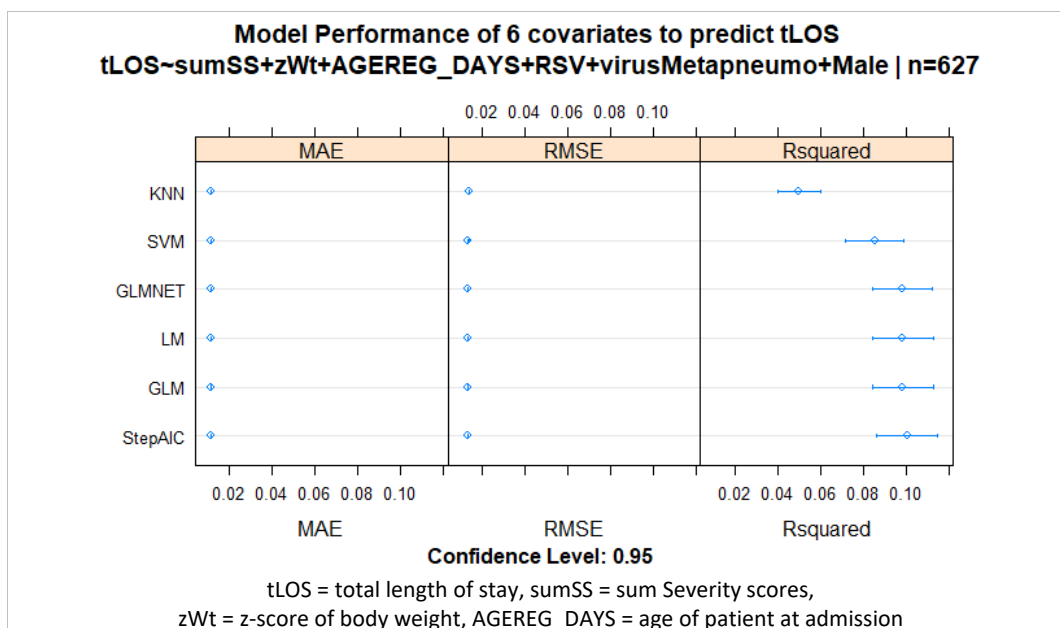
Fig. 25: Overview Model Performance for six covariates to predict total Length of Stay (adjusted values)

5.4.3   Severity score of the first/second six hours of the hospital stay vs. tLOS

In this part of the analysis, linear regression was used to explore the relationship between the cumulative **severity score** of the **first six hours** and the **second six hours** of the hospital stay. A simple scatterplot over all cases with sufficient data, as shown in Fig. 26, suggested low correlation between the cumulative severity score of the second six hours of hospital admission and length of stay. As discussed above, the total length of stay was transformed using the Yeo-Johnson Power transformation, and the severity score log-transformed. The results of linear regression for the cumulative score of the second six hours were significant.

```
Call:
lm(formula = mFormula, data = md)

Residuals:
      Min         1Q      Median         3Q         Max
-0.0186251  -0.0073911  -0.0001273   0.0065284   0.0194119

Coefficients:
```

```
                    Estimate Std. Error  t value Pr(>|t|)
(Intercept)        2.235e+00  1.057e-03 2114.522   <2e-16 ***
sumSS_1st_6hours -9.662e-05  3.537e-04   -0.273   0.7849
sumSS_2nd_6hours  9.608e-04  3.736e-04    2.572   0.0104 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.008461 on 496 degrees of freedom
Multiple R-squared:  0.01698,      Adjusted R-squared:  0.01302
F-statistic: 4.284 on 2 and 496 DF,  p-value: 0.0143
```

Code 8: Summary of Generalized Linear Model for the Severity Score of first/second six hours vs. Length of Stay
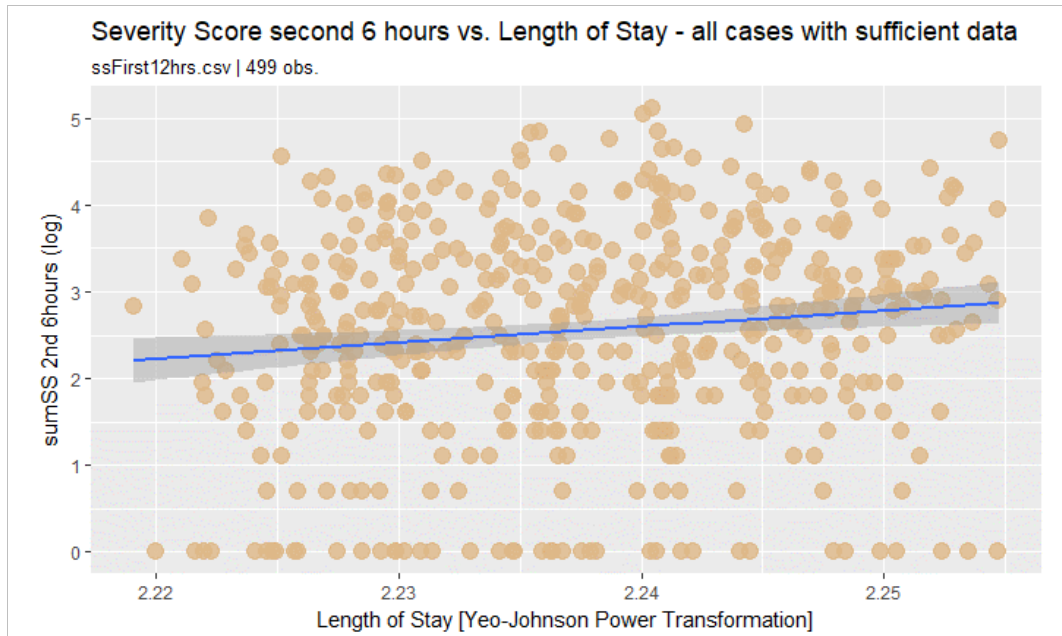
Fig. 26: Cumulative Severity Score of second six hours vs. Length of Stay (Yeo-Johnson transformed)

5.5 Odds Ratios for Prolonged Hospital Stay

To test the odds ratios of nine variables in relation to prolonged hospital stay greater than five days, the R package *oddsratio* was used to calculate odds ratios from a generalized linear model (Code 9).[89] The increments for numerical variables were as follows: severity score = 20, z score of body weight = 0.5, and for admission age in days = 30. The median length of stay was 3.93 days.

Complete code online: http://ckcdata.com/R-code/5.5-Correlation-SeverityScore-Length-of-Stay.html

```
## Summary tLOS [days]
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  2.032   2.981   3.926   4.446   5.565  10.314
...

fit.glm <- glm(prolongedLOS ~ ., data=md_, family = binomial)
or.glm <- or_glm(data = md_, model = fit.glm,
           incr = list(SeverityScore=20, zScore_BodyWeight=0.5, Age_Days=30))
…
```

Predictor	Oddsratio	CI_low (2.5 %)	CI_high (97.5 %)	increment	glm – Level of Significance
SeverityScore	1.215	1.024	1.444	20	*
Male	1.193	0.811	1.764	Indicator variable	
SimpleCase	0.403	0.265	0.606	Indicator variable	***
Prematurity	1.174	0.520	2.560	Indicator variable	
zScore_BodyWeight	0.875	0.816	0.936	0.5	***
Age_Days	0.922	0.866	0.980	30	**
RSV detected	1.450	0.957	2.210	Indicator variable	.
virusRhinoEntero detected	0.898	0.347	2.137	Indicator variable	
virusMetapneumo detected	2.322	1.065	4.995	Indicator variable	*

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Code 9: Calculate Odds Ratios Using Generalized Linear Model
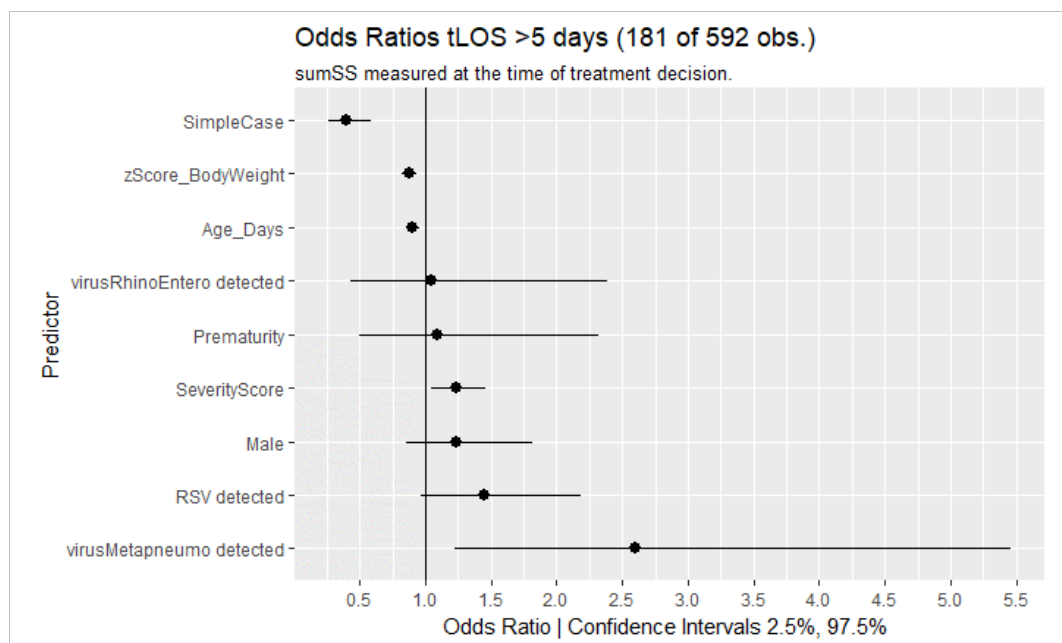
Fig. 27 Odds Ratios for Prolonged Hospital Stay (>5 days)

The results (Code 9, Fig. 27) show that simple cases, z score of body weight and admission age have a negative relationship with length of stay greater than five days. For example, the odds for a simple case (acute viral bronchiolitis as the only diagnosis) to stay in hospital longer than five days is reduced by approx. 50%. It is approx. 10% reduced for z score of body weight and

admission age. The odds ratios for severity score and Metapneumovirus show a positive relationship which is significant because the confidence interval does not contain 1. The odds ratio for the severity score was 1.25 and for Metapneumovirus 2.32. The absolute and relative numbers of RSV and Metapneumo virus are depicted in Table 20. The overall occurrence of Metapneumo virus was much lower than RSV.

Table 20: Absolute and Relative Occurrence of RSV And Metapneumo Virus

Description	RSV	Metapneumo Virus	Co-infection RSV and Metpneumo virus
LOS <=5 days	198 (66%)	21 (58%)	
LOS > 5 days	89 (31%)	15 (42%)	1
Total	287	36	
Percentage of all cases (n=574)	50%	6.3%	

5.6 Discussion

5.6.1 Findings

The focus of this chapter was prediction of length of hospital stay. It is important to note that the dataset of this retrospective, observational study was restricted to patients with acute viral bronchiolitis who required either low-flow (NC) or high-flow (HFNC) oxygen applied via nasal cannula (inclusion criteria). Patients without respiratory support were not included in the analysis.

Feature selection, performed with five different methods, identified the most important predictors of length of stay. Corrected age at the time of admission and z score of body weight demonstrated the strongest influence. With regards to the influence of the viral cause on length of stay, Metapneumo virus ranked higher than RSV. However, RSV is clinically more relevant because of its higher incidence. RSV accounted for 50% of 574 cases, and Metapneumo virus only 6.3% (Table 20). When investigating individual parameters of respiratory distress, capillary pO_2, signs of retractions and respiratory effort, breath sounds, heart rate, body temperature, respiratory rate, and pulse-oximetry were among the highest-ranking parameters.

Under the assumption that disease severity is best described by the sum of individual, weighted scores, as described in chapter 4 (p. 55), the resulting severity score (*sumSS*) was used for correlation studies. The relationship of the respiratory severity score with length of stay was examined. A re-run of the feature selection algorithms with the respiratory severity score and five other variables, as described in Table 18, demonstrated a ranking of the severity score almost as high as age and z score of body weight.

To test the strength of the severity score to predict hospital length of stay in conjunction with five other variables, ten-fold cross-validation, repeated 10 times, was applied to six models. The severity score measured at the <u>time of the treatment decision</u> to apply either standard nasal oxygen or high flow nasal oxygen was highly significant in predicting total length of stay (p<0.001). The cumulative severity score, calculated from data obtained during the <u>second six-hour period</u>

following the first recorded event, was also significant (p<0.01), albeit not as high. Finally, odds ratio calculations were used to predict hospital stay greater than five days. Cases with acute viral bronchiolitis as the only diagnosis (*simpleCase*), z score of the body weight, and age showed a significant negative relationship with prolonged length of stay. Severity score, measured at the time of the treatment decision, demonstrated a significant positive relationship.

Previous studies have investigated the influence of predictors on hospital admission,[39,90-92] disease severity,[69] ICU admission,[93] and prolonged hospital stay.[94] Length of stay is often used as the dependent, primary outcome variable in intervention studies, which will be discussed later in chapter 6 and 7.

Age, low birth weight and prematurity are considered important predictors of severe bronchiolitis[48] and prolonged hospital stay.[94,95] Only a few studies have attempted to predict prolonged length of stay with these variables.

Age

Most recently, Rodriguez-Martinez and colleagues defined prolonged hospital stay as greater than five days. For age, they found an odds ratio of 0.92 (95% CI 0.84 to 0.99; p=0.049) after adjusting for comorbidities and other patient parameters.[94] Jartti and colleagues produced very similar results when investigating the influence of young age (<2 months) on hospital length of stay of equal or more than three days: OR 2.70 – 95% CI 1.08–6.73; p=0.03.[96]

The current study, which did not apply any adjustments, confirms these results. The odds ratio of age was 0.92 (97.5% CI 0.87 to 0.98; p<0.01). The calculation was based on a generalized linear model which incorporated eight other covariates (Fig. 27, p. 68). Age-specific calculations revealed an odds ratio of 1.88 (97.5% CI 1.24 to 2.86; p=0.003) for infants less than two months old. The odds ratio for infants aged two to 6 months did not show any significant relationship with length of stay.

Z Score Body Weight

Prematurity and low birth weight increase the risk for severe disease including invasive and non-invasive respiratory support, and consequently increased length of stay. [48,94,96,97]

The current dataset did not provide enough data about birth weight and prematurity. As an alternative, the corrected age and the median body weight obtained during the hospital stay were used to calculate the z score of the body weight. The z score between minus two and plus two indicates normal body weight. To the author's knowledge, z score calculations have never been used before to assess age-specific (if necessary, corrected for gestational age) and gender-specific weight gain. The overall odds ratio of the z score of the body weight was 0.91 (97.5% CI 0.85 to 0.96; p=0.001), indicating a negative relationship. The odds ratio for a patient with a z score of the

body weight less than minus two was 2.45 (97.5% CI 1.35 to 4.45; p=0.003). Whereas a normal z score between minus two and plus two was not significantly related with length of stay.

RSV, Metapneumovirus

RSV is the most common cause of acute viral bronchiolitis. Several studies have compared the course of disease in relation to the presence or absence of RSV. Hervás and colleagues reported that infants infected with RSV had longer median hospital stay than patients infected with other viral causes (RSV six days vs. non-RSV five days).[95] Jartti and colleagues found that bronchiolitis patients infected with rhinovirus only had a shorter hospital length of stay than infants with RSV only bronchiolitis. They found that the odds ratio of Rhinovirus only was 0.45 (95% CI 0.22–0.92; p=0.03). The current study also identified RSV as the most common viral cause (50%), however, calculation of the odds ratio did not reveal a significant relationship with length of stay. As a new finding, infection with Metapneumo virus was significantly related with length of stay. The odds ratio was 2.32 (97.5% CI 1,07 to 5; p=0.024). This result must be interpreted on the background of the low percentage of Metapneumo virus of only 6.3% (Table 20, p.69).

Severity score

The odds ratio of the severity score measured at the time of the treatment decision NC vs. HFNC for prolonged hospital length of stay was 1.21 (97.5% CI 1.03 to 1.43; p=0.023), indicating a positive relationship. To the author's knowledge, there is only one recent study that used a severity score to predict hospital length of stay longer than three days. Golan-Tripto and colleagues applied the modified Tal score (MTS) on admission. They concluded that the MTS could "predict fairly LOS". The results did not consistently reach statistical significance.[69]

Predictive Modeling

Rather than investigating clinical parameters in isolation, one of the objectives of this study was to develop and test machine learning models that incorporated the most important covariates for predicting hospital length of stay.[98] Feature selection helped to determine the most important variables. In the next step, the data was evaluated and prepared, as described in chapter 5.4 (p. 61). The R package *caret* was used to apply ten-fold cross-validation, repeated 10 times, to assess the performance of six different models. The model with the best performance indicators, such as R^2, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), was a generalized linear model with stepwise feature selection (Code 7, p.66). The following predictors demonstrated highly significant correlations with total length of stay (tLOS): severity score measured at the time of the treatment decision to apply NC or HFNC, z score of body weight, age expressed in days, detection of RSV, and detection of Metapneumo virus.

Table 21: Performance Parameters for Generalized Linear Model with Stepwise Feature Selection to predict length of stay

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
MAE	0.0099	0.0108	0.0113	0.0112	0.0116	0.0126	0
RMSE	0.0118	0.0129	0.0134	0.0133	0.0138	0.0148	0
Rsquared	0.0012	0.0469	0.0817	0.1007	0.1498	0.3113	0

5.6.2   Limitations

The current study has several limitations. The retrospective, single-centre design of the analysis was based on observational data which was obtained mainly for documentation and billing purposes. The dataset was incomplete with regards to information about prematurity and birth weight. The datapoints for calculation of the severity score were not systematically and consistently recorded. This study did not investigate the effect of medications, e.g. beta-sympathomimetics, epinephrine, corticosteroids, antibiotics, or other therapeutic interventions, such as inhalation of hypertonic saline. The focus was on application of non-invasive respiratory support. Due to inconsistent and incomplete data, the true length of hospital length of stay could not be determined. Ideally, readiness for discharge should have been evaluated by applying standardized discharge criteria.[73]

5.6.3   Conclusion

The current study introduced a new way of calculating a respiratory severity score which demonstrated a significant positive correlation with hospital length of stay. It's odds ratio for predicting prolonged length of stay (>5 days) was elevated and statistically significant ($p<0.05$). In addition, z scores of the body weight were introduced as highly significant predictor of length of stay.

Predictive modeling showed that a generalized linear model with stepwise feature selection produced low mean absolute error (MAE) and low root mean squared error (RMSE) when using severity score in conjunction with five other variables. Even though there were significant correlation coefficients, the low R-squared value of 0.1 indicated that only 10% of the variability in the dependent variable was explained by the model.

Contrary to current literature, the influence of RSV on prolonged hospital stay did not reach significance. The odds ratio of Metapneumo virus was significantly elevated. However, this was based on small numbers.

5.6.4   Consideration for Future Research

A prospective study design with consistent and more accurate documentation of the clinical parameters of the respiratory severity score might help to build more accurate models for predicting hospital length of stay.

6 Prediction of High Flow Therapy

6.1 Feature Selection

In this chapter, seven different methods of feature selection were used to determine the most important predictors of high flow therapy. The summary of the results is listed in Table 22.

Table 22: Results of Feature Selection for HFNC

	VarName	glm	LASSO AUC	BORUTA rank	LM StepAIC	BART rank	XGBOOST rank	GBM	Count Selected
1	ssRespRetraction	*	3	2	***	2	1	1	7
2	ssRespEffort	**	1	1	**	1	2	2	7
3	ssRoomAir	***	4	3	***	3	3	3	7
4	ssCBG_pCO2	*	2	8	***	4	6	9	7
5	ssRespPEWS	**		9	**	5	5	4	6
6	ssConscious			6	***	7	10	7	5
7	ssRespType		5	5	*	9			4
8	ssNasalFlaring			4		10	8	5	4
9	virusMetapneumo	***			***	6			3
10	AGEREG_DAYS						4	6	2
11	zWt						7	10	2
12	ssRR					8	9		2
13	ssUpperAirwaySounds			7	*				2
14	RSV				*				1
15	virusRhinoEntero				*				1
16	ssHR							8	1

Fig. 28 and Fig. 29 depict the ranks of the features as calculated by the BART algorithm and extreme gradient boosting, respectively. The complete R script can be found online: 6.2-Feature Selection-HFNC-v2.3.html.
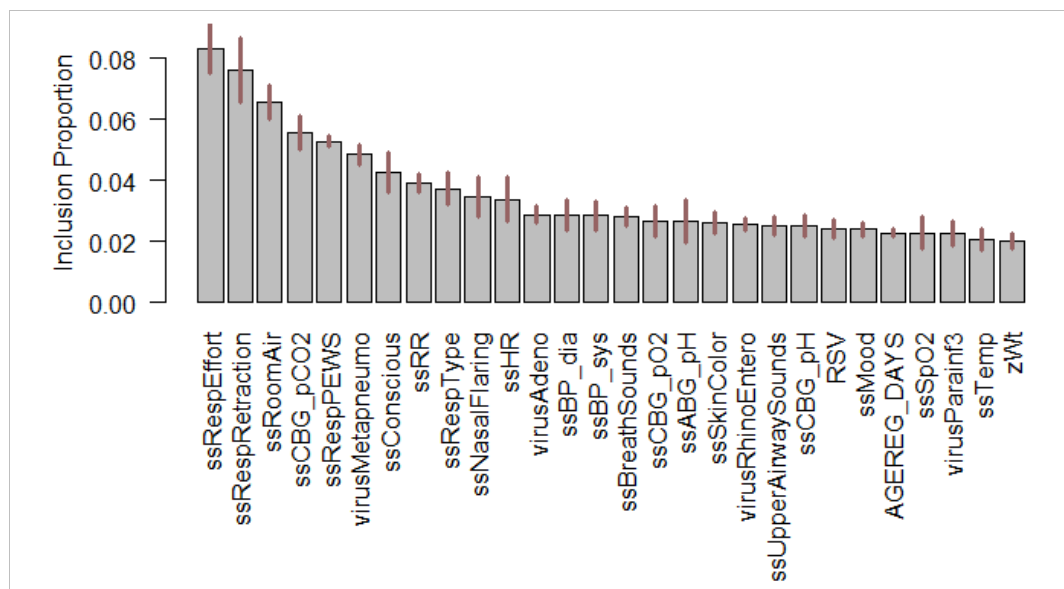
Fig. 28: Feature Selection High Flow Therapy: `bartMachine` – *investigate_var_importance*
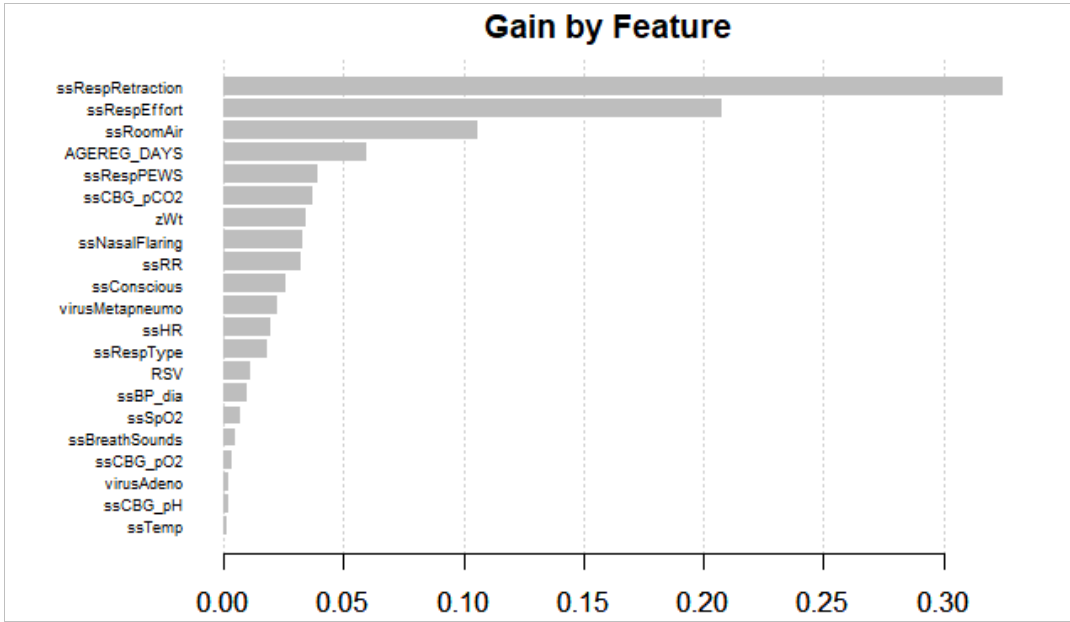
Fig. 29: Feature Selection High Flow Therapy; Extreme Gradient Boosting (XGBOOST)

The first nine individual scores with the highest relative influence on the commencement of high flow therapy were (approximately) ranked in this order: ssRespEffort, ssRespRetraction, ssRoomAir, ssCBG_pCO2, ssRespPEWS, ssConscious, ssRespType, ssNasalFlaring, virusMetapneumo (Table 22).

Some of the most important features were clinical observations that described symptoms of respiratory distress. Table 30 (page 118) provides a comprehensive list of the words found in the current dataset, and their corresponding weights assigned by the investigators. As discussed in chapters 4.1 (p. 43) and 5.2.7 (p. 58), the design and structure of these categorical items were not optimised for analysis. As an example, Table 23 demonstrates how "Head Bobbing" was defined multiple times. This indicates that the current design of the database showed reduce usefulness for a reliable analysis of individual items.

Table 23: "Head Bobbing" as an Example of Multiple Use of keywords

ssRespRetraction	All Muscles Used=4, head bob=3, **head bobbing=3**, Intercostal=2, Mild=2, Moderate=4, **Other: Head bobbing=3**, **Other: Head bobbing.=3**, Severe=6, Subcostal=1, Substernal=1, Supraclavicular=3, Suprasternal=3
ssRespEffort	Increased=1, Mild Distress=2, Moderate Distress=3, Other: Grunting=4, **Other: Head bobbing=3**, Retractions=1, Severe Distress=4, tachypneic=1
ssRespEffort (ED)	Grunting=3, Nasal Flaring=2, **Other: head bobbing=3**, Other: tachypneic=1, Retractions=2
ssRespType	Agonal=4, Apnea=3, Bradypnea=2, Controlled with Ventilator=5, Dyspnea=2, Gasping=4, Grunting=3, HFOV or Jet Vent - Unable to Assess=4, increased retractions=2, increased work of breathing=2, Irregular=1, nasal flaring=2, **Other: head bobbing=3**, Other: WITH NASAL CANNULA=2, Shallow=1, subcostal retraction=1, Tachypnea=2, With Ventilator=5

Due to redundancies and the retrospective nature of this study, the author decided not to further evaluate and calibrate the individual scores. Instead, the new respiratory severity score, summing all 22 individual scores, was regarded as a practical work-around for the purpose of this retrospective analysis.

6.2 Predictive Value of 6- and 12-hour Severity Score for High Flow Therapy

As described in chapter 0 (p. 50), the sums of the weighted, individual scores were calculated for the first and second six hours of the total hospital stay to determine their ability to predict the need for high flow therapy. Logistic regression (GLM) was significant for both time periods (0-6 hrs $p<0.001$, 6-12 hrs $p<0.001$). The predictive value of both severity scores was evaluated by testing eight models (GLM, LDA, GLMNET, KNN, CART, Naïve Bayes, SVM, glmStepAIC) with ten-fold cross-validation, repeated ten times (Fig. 30).
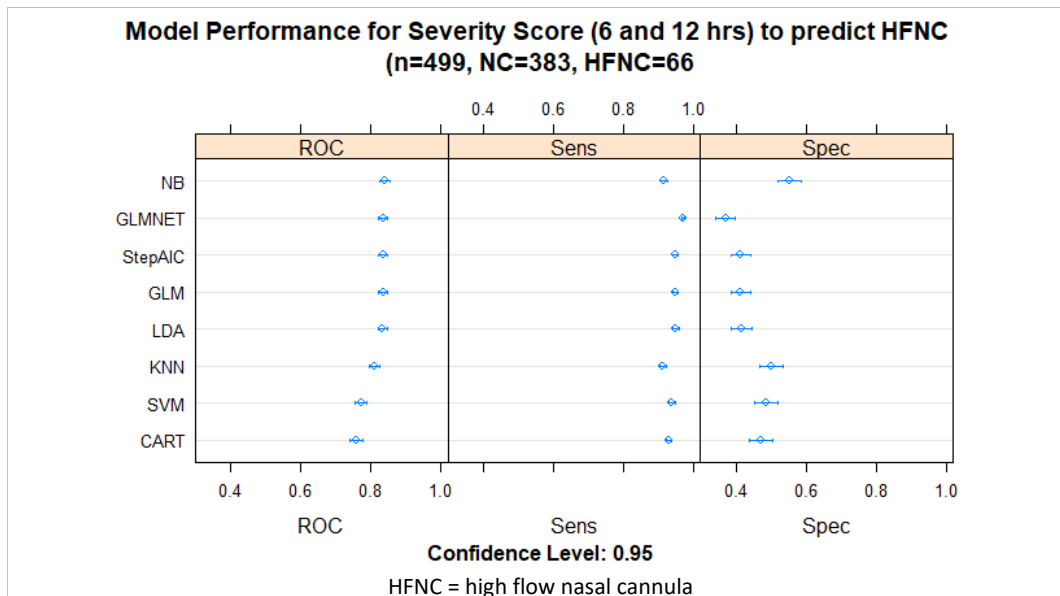
Fig. 30: Model Performance for 12h Severity Score to predict High Flow Therapy

Overall, the models calculated a ROC between 0.76 and 0.84, the sensitivity was above 0.91. The highest specificity (true negative rate) was 0.55 (NB = Naïve Bayes). In this context, it would mean that 55% of the patients who do not require HFNC would correctly be identified. The ROC curve for the Naïve Bayes model is shown in Fig. 31.
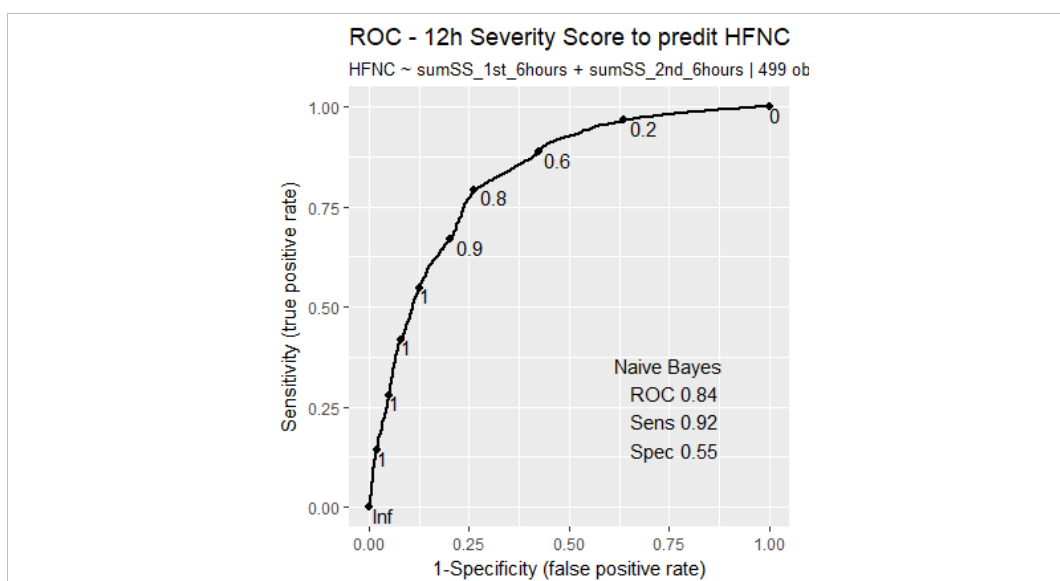
Fig. 31: ROC for 12h Severity Score to predict the High Flow Therapy

Odds ratio calculations revealed significant results for both time periods (Fig. 32). Odds ratios of the severity score for the first six hours was 1.39 (97.5% CI 1.23 – 1.59) and for the second six hours 2.2 (97.5% CI 1.75 – 2.81).
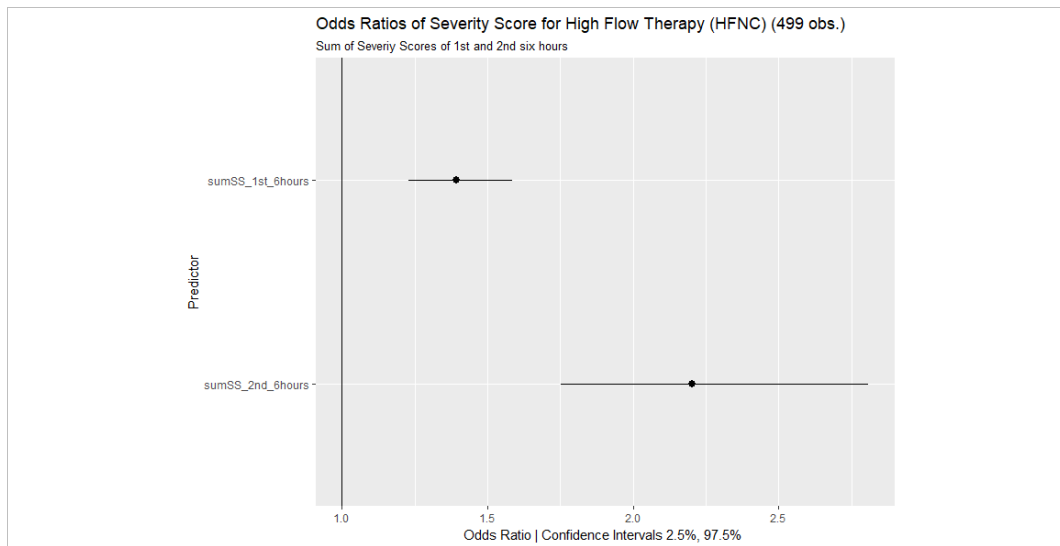
Fig. 32: Odds ratios for severity scores first and second six hours to predict High Flow Therapy

6.3 Predictive Modeling for Predicting High Flow Therapy

This part of the analysis looked at the severity score measured at the time of the treatment decision to apply either standard nasal oxygen or high flow nasal oxygen. In the following, the R package *caret* was used to test the performance of seven models by applying ten-fold cross-validation, repeated ten times. As discussed in chapter 5.4 (p. 61), Yeo-Johnson Power transformation was applied to the z score of the body weight and to age. Log transformation was applied to the severity score. The combination of severity score with six other covariates produced a high predictive value with regards to commencement of high flow therapy.
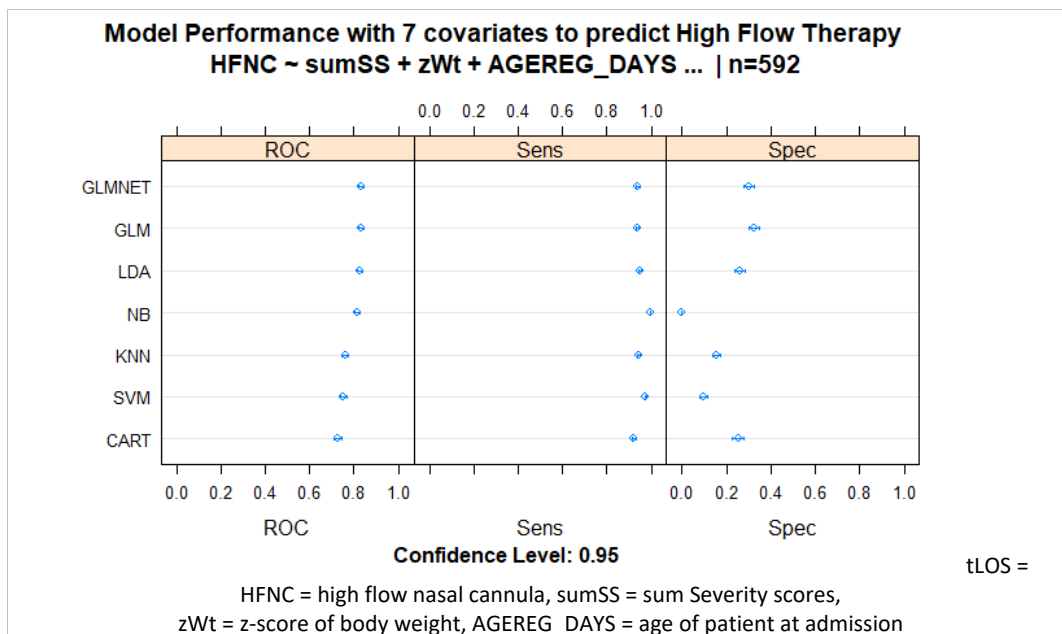
Fig. 33: Performance Overview of seven Models for 7 Covariates Predicting the Use of High Flow Oxygen Therapy

The GLM model demonstrated the best performance with a high area under the ROC curve of 0.83, sensitivity of 0.94, and specificity of 0.33 (Fig. 34).
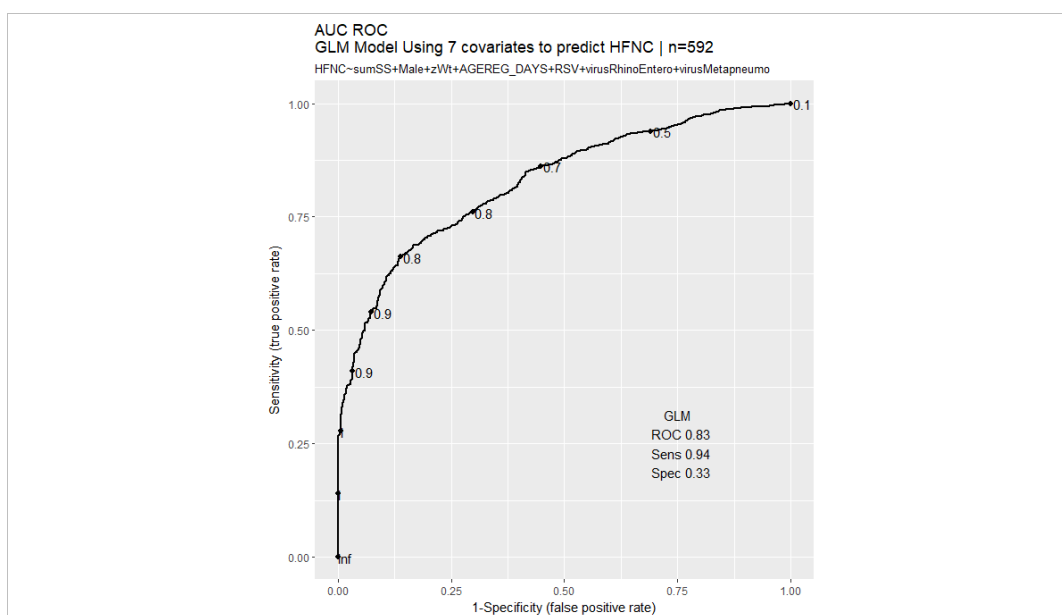
Fig. 34: Performance of GLM Model using 7 Covariates to Predict High Flow Oxygen Therapy

6.4 Odds Ratios for High Flow Therapy

Calculation of the odds ratios of seven covariates to predict high flow therapy revealed significant results for respiratory severity score, age younger than 100 days, z score of the body weight, and the presence of RSV, RhinoEntero virus and Metapneumo virus (Table 24, Fig. 35).

Table 24: Odds ratios to predict High Flow Therapy

Predictor	oddsratio	CI_low (2.5 %)	CI_high (97.5 %)	increment
SeverityScore	2.777	2.227	3.519	20
Age <100 days	2.965	1.737	5.190	

SimpleCase	0.740	0.453	1.197	
zScoreBodyweight < -2	2.568	1.174	5.470	
RSV detected	2.190	1.284	3.825	
virusRhinoEntero detected	4.060	1.688	9.829	
virusMetapneumo detected	6.398	2.622	15.625	
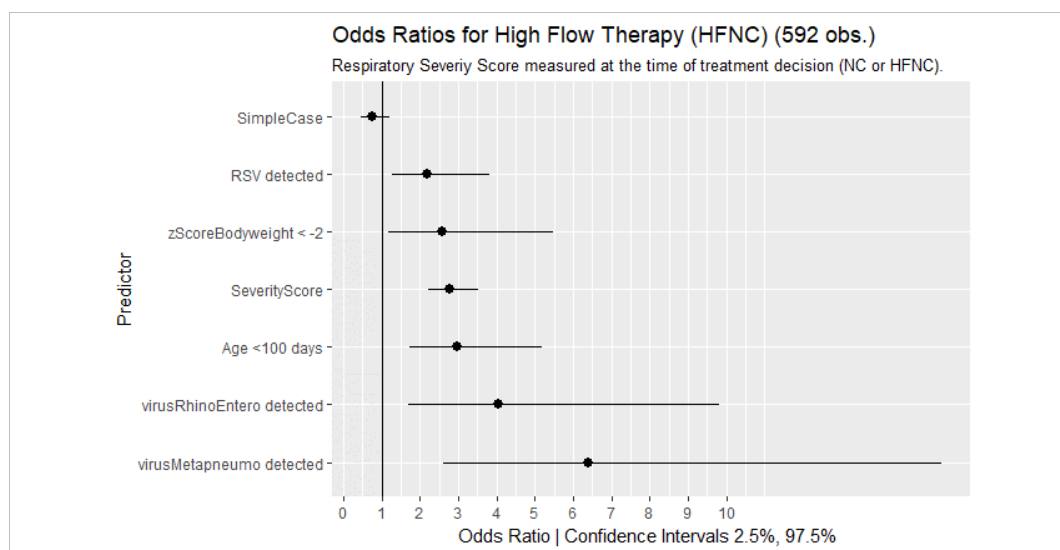
Fig. 35: Odds Ratio of Seven Covariates for prediction of High Flow Therapy (HFNC)

6.5 Discussion

The most important result of this chapter is the finding that a data-driven respiratory severity score is significantly related to high flow oxygen therapy. Examination of the severity score obtained from three different time windows (time periods 1,2, and 3; see also chapter 4.3, p. 50) yielded a significant relationship with the commencement of high flow therapy.

The severity score obtained in the second six hours of the total stay (emergency department and/or hospital) had the highest ROC AUC (0.84) when using a Naïve Bayes model.

When using the severity score obtained at the time of treatment decision, in conjunction with six other covariates, the generalized linear model was the best fitting model to predict high flow therapy. The ROC AU was 0.83 (sensitivity 0.94, specificity 0.33). The odds ratio of the severity score was 2.8 (97.5% CI, 2.23 to 3.52). In addition, Z score of the body weight smaller than minus two, age younger than 100 days, and RSV, RhinoEntero virus and Metapneumo virus also had a significant positive relationship with commencement of high flow therapy.

Several studies have investigated predictors of treatment escalation, such as admission to paediatric intensive care[99,100], commencement of continuous positive airway pressure (CPAP)[45], or conventional mechanical ventilation (CMV).[48,100,101] Many studies have compared the rate of invasive ventilation before and after the introduction of methods of non-invasive ventilation (low flow, high flow, CPAP).[44,47,102,103] However, to the author's best knowledge, there are no publications on predictors for high flow therapy. Only two studies have investigated indications

for high flow therapy or CPAP. Long and colleagues found that he most common indication for high flow therapy was bronchiolitis (69%). Indications for high flow therapy were broadly defined as moderate to severe respiratory distress where increased work of breathing or hypoxaemia was not relieved by standard oxygen therapy. The following clinical parameters were recorded on admission and two hours after initiation of high flow therapy: were respiratory rate, heart rate, body temperature, and oxygen saturation. However, no attempt was made to analyse these vital signs or formally assess severity of illness.[104] Evans and colleagues found in a small retrospective study that the variables with the strongest negative or positive relationship with nasal CPAP (n=28) were: supplemental oxygen in emergency department, oxygen saturation, age, respiratory rate, heart rate, Glasgow coma scale, and gestational age.[45]

The big advantage of machine learning is testing a multitude of predictors at once. This is more likely to reflect reality more correctly and advance our knowledge on how to make data-driven therapeutic decisions. Instead of focusing on individual factors, the aim of the current study was to build and test prediction models that were based on several covariates like respiratory severity score and other important covariates, as identified by feature selection. The generalized linear model showed the highest performance (ROC AUC 0.83) in predicting high flow therapy.

6.5.1   Limitations

The main limitation of this study was its retrospective and single-centre design. Therefore, data entry was not standardised and not evaluated. Some variables, deemed important for severity assessment of acute viral bronchiolitis, were not available, for example, oxygen saturation in room air, oxygen requirement (FiO_2) to maintain SpO_2 greater than 92%. Another limitation was the lack of the exact date and time of start and end of non-invasive ventilation. Often this information had to be indirectly obtained by manual assessment of the patient's clinical chart.

Finally, the use of z scores for heart rate and respiratory rate could have potentially added more accuracy. At the time of the analysis, these z score values were not available.

6.5.2   Future Research

Further studies are required to address issues mentioned under limitations. Prospective studies are needed to assess how real-time application of a data-driven model (GLM) supports the clinical decision-making process. Useful outcome variables should be success rate of various respiratory support treatments such as low flow therapy, high flow therapy, and nasal continuous positive airway pressure.

7 Comparative Effectiveness of High Flow and Standard Therapy

This chapter investigates the treatment effect of high flow therapy on hospital length of stay. Feature selection and propensity score matching methods were used to identify the propensity score balance ("common support") and find the best possible matching between control group and treatment group. Several methods of matching were applied ("optimal", "greedy", "full", "genetic"). The aim of propensity score matching was to eliminate indication bias.

7.1 Description of Dataset Used for Propensity Score Matching

To answer the question whether any use of high flow therapy influenced total length of stay, cases were marked with either *HFNC=1* identifying high flow therapy with or without escalation, or *HFNC=0* representing standard therapy (=low flow oxygen via nasal cannula) without escalation. The resulting dataset comprised 499 observations in total (standard therapy: 383 successful cases; high flow therapy: 98 successful cases and 18 failed cases).

7.2 Covariate Selection

Chapter 6 (p. 73) discusses which variables had the strongest influence on treatment assignment, i.e. high flow therapy. Because of limitations of the current dataset, the sum of weighted, individual scores was used rather than individual scores. Table 24 (p. 77) lists the odds ratios of the most important variables for prediction of high flow therapy. Respiratory severity scores from two different time windows were used (time period 2 and 3) for two separate runs of the R script. The complete output of the underlying R scripts can be found online: http://ckcdata.com/R-code/.

As before, the Yeo-Johnson Power transformation was applied to total length of stay, z score of body weight, and age. Log transformation was applied to severity score.

7.3 Estimate of the Propensity Model

The R function *glm* was applied to fit a generalized linear model using logistic regression ("family=binomial") with the following formula (Code 10) to estimate propensity scores. The propensity score is the probability of treatment assignment conditional on the observed baseline covariates.[105]

```
psFormula <- HFNC ~ Male + AgeDays + zWt + Simple + sumSS (or sumSS_2nd_6hours) + RSV + virusRhin
oEntero + virusMetapneumo
glm1  <- glm(psFormula, family=binomial, data=md)
md$glm1 <- log(fitted(glm1)/(1-fitted(glm1)))
md$glm1_fitted <- glm1$fitted
```

Code 10: R Formula for Estimating Propensity Scores

The back-to back histogram of the propensity score distribution for the control and treatment group is shown in Fig. 36. There appears to be enough overlap between both groups to perform advanced methods of matching.
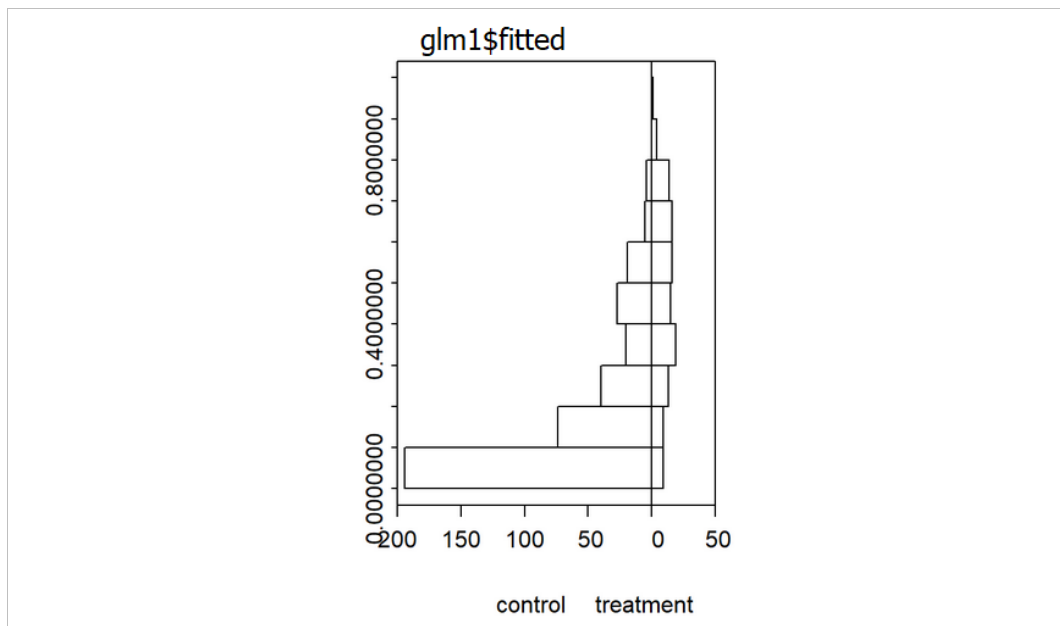
Fig. 36: Distribution of Propensity Scores for Control and Treatment Group

Other methods of propensity score estimation can be used to plot the propensity score distribution depicting the balance on the propensity scores (common support).

```
param <- matchit(mFormula, data = md)
md$param ps <- param$distance

set.seed(123456)
nonparam <- ps(psFormula, data = md, n.trees = 5000, interaction.depth = 4,
          shrinkage = 0.01, stop.method = "es.mean", estimand = "ATT")
md$nonparm_ps <- nonparam$ps$es.mean
```

Code 11: Estimating propensity score by logistic regression and boosted regression
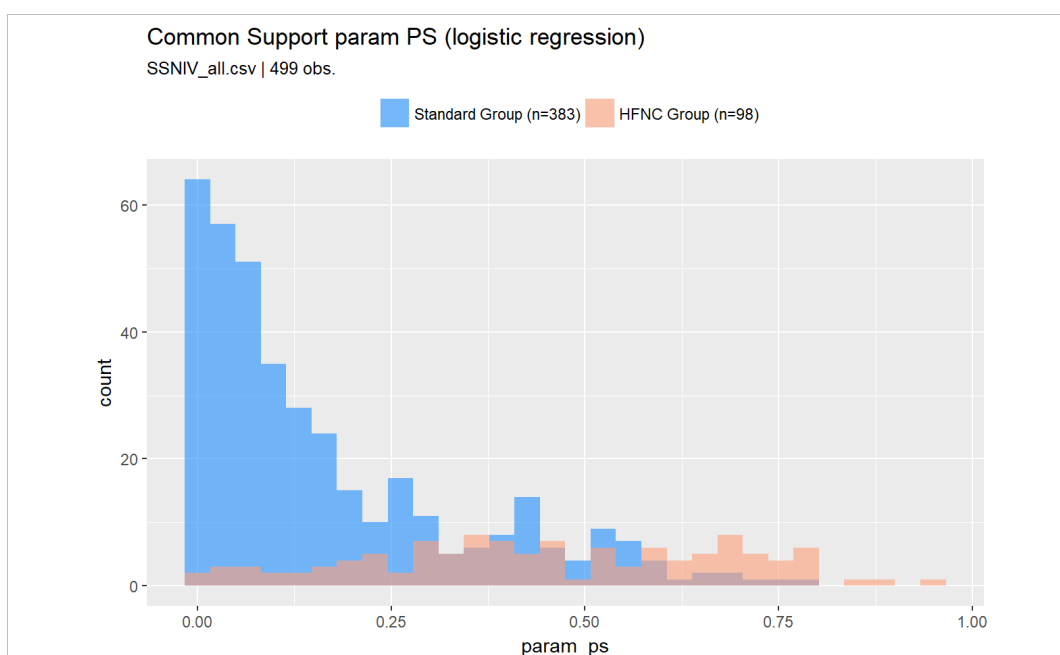
Fig. 37: Common Support of Propensity Scores Using Logistic Regression
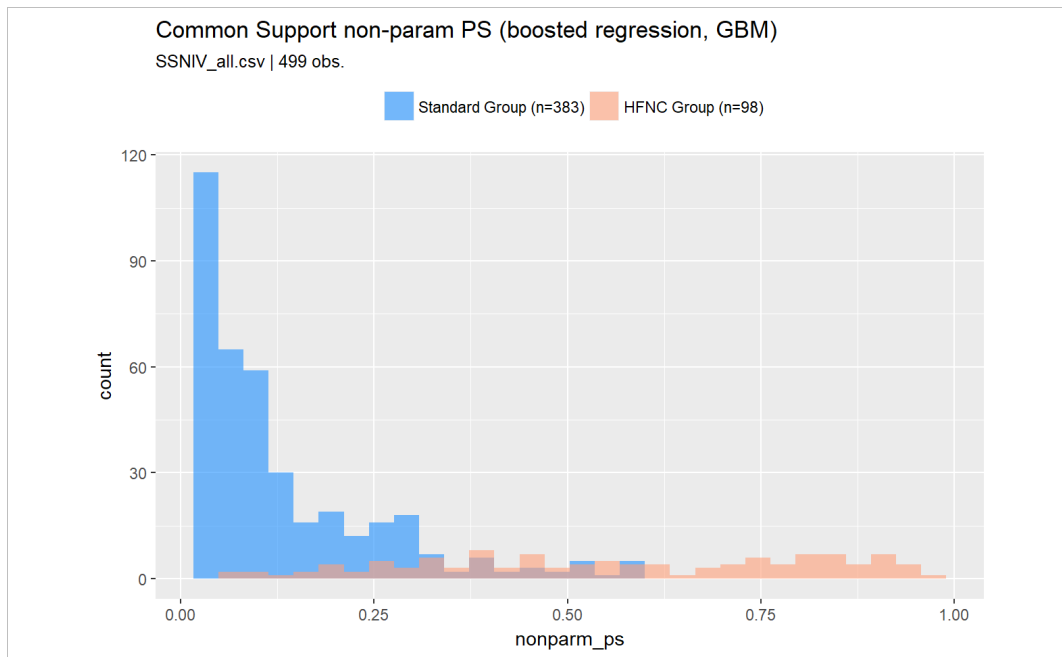
Fig. 38: Common Support of Propensity Scores Using Boosted Regression

7.4 Propensity Score Matching

There are two R packages available that perform propensity score matching with balance optimization of confounding variables: Matching[106] and MatchIt[107]. The current study has incorporated both packages in one R script (available online: http://ckcdata.com/R-code). Development of the script was based on practice guides which can be found in the latest literature on this topic.[29,108,109]

Table 25: Results of Different Methods of Covariates Matching – R Package *Matching*

R Package Matching							
R U N	**Method**	**Matched Observations Control/Treated**	***MatchBalance* Variables**	**T-test p-value**	**Estimate tLOS (raw) ATT**	**p.val (transformed)**	**Sign. Level**
1	Match	116/116	glm1_fitted	0.063329	920	0.023071	*
2	Match	116/116	glm1_fitted Male AgeDays zWt Simple sumSS_2nd_6hours RSV virusRhinoEntero virusMetapneumo	2.6204e-10 0.024093 0.094828 0.88193 0.044109 2.9503e-07 0.15641 1 0.31732	647	0.049128	*
3	Match	116/116	Male AgeDays zWt Simple sumSS_2nd_6hours RSV virusRhinoEntero virusMetapneumo	0.04410 0.16978 0.99036 0.15641 9.5589e-10 0.15641 0.31732 1	919	0.0055415	**
4	Genetic Matching	116/116	Male AgeDays zWt Simple sumSS_2nd_6hours RSV virusRhinoEntero virusMetapneumo	0.044109 0.16978 0.99036 0.15641 9.5589e-10 0.15641 0.31732 1	919	0.0055415	**

Table 25 summarizes the results of the R function *MatchBalance*. All four runs produce low p-values for high flow therapy to have a significant positive effect on total length of stay (tLOS). The estimated average treatment effect on the treated patient (ATT) is shown in column 4. For easy understanding, the total length of stay expressed in minutes is used. The actual calculations of significance were based on transformed values.

These results seem to indicate that the use of high flow therapy prolonged hospital length of stay by ten to fifteen hours, i.e. 600 to 900 minutes. However, the p-values for the T-test comparing the difference between control group and treatment group remained significantly low for gender (Male), respiratory severity score (sumSS_2nd_6hours), and in run number two also for simple cases (Simple) and propensity scores (glm1_fitted). Therefore, incomplete matching between treatment group and control group potentially accounted for the difference in the treatment effect on outcome (LOS).

Table 26: Results of Different Methods of Covariate Matching – R Package *MatchIt*

R Package MatchIt							
RUN	Method	Matched Observations Control/Treated	Summary of Matching Variables	Std. Mean Diff.	Estimate tLOS (raw) ATT	p.val (transformed)	Sign. Level
5	Greedy Matching	94/94	AgeDays	0.0949	400	0.15190	NS
			distance	0.0772			
			Male	0			
			RSV	0.0652			
			Simple	0.0432			
			sumSS_2nd_6hours	0.0879			
			virusMetapneumo	0			
			virusRhinoEntero	0			
			zWt	-0.0317			
6	Optimal Matching	116/116	AgeDays	0.1201	638	0.0277	*
			distance	0.358			
			Male	0.0877			
			RSV	-0.0352			
			Simple	-0.1049			
			sumSS_2nd_6hours	0.2802			
			virusMetapneumo	0.0272			
			virusRhinoEntero	0.079			
			zWt	-0.0211			
7	1:2 Matching	144/94	AgeDays	0.0842	650	0.02963	*
			distance	0.0988			
			Male	0.0433			
			RSV	-0.0109			
			Simple	0.0216			
			sumSS_2nd_6hours	0.1094			
			virusMetapneumo	0			
			virusRhinoEntero	0.0325			
			zWt	0.0149			
8	Full Matching	383/116	AgeDays	0.026	800	0.006814	**
			distance	0.1088			
			Male	0.0096			
			RSV	0.0014			
			Simple	0.0008			
			sumSS_2nd_6hours	0.1382			
			virusMetapneumo	0.0172			
			virusRhinoEntero	0.018			
			zWt	0.0218			
9	Genetic Matching	74/116	AgeDays	0.0083	1096	0.00222	**
			distance	0.0898			
			Male	0.0877			
			RSV	-0.1057			
			Simple	0.1224			
			sumSS_2nd_6hours	0.0879			
			virusMetapneumo	0			
			virusRhinoEntero	0.1054			
			zWt	0.1188			

http://ckcdata.com/R-code/7.2-Propensity-Score-Matching-MatchIT-v1.2.html

Table 26 shows the average treatment effect for the treated (ATT) as estimated by various matching methods (run 5-9) provided by the R package *MatchIt*. Genetic matching produced the lowest p-value (0.0022) based on 74 controls versus all 116 high flow cases. Low values of standardized mean differences (Std Mean Diff.) indicated a satisfactory matching result. The estimated average treatment effect was prolongation of total length of stay by approx. 1000 min.

The following figures demonstrate some of the results of matching. The jitter plot in Fig. 39 compares the distribution of the propensity scores for genetic and greedy matching. In the case of greedy matching, the algorithm eliminated 22 treatment units with high propensity scores and 289

control units with low propensity scores. The resulting match between control and treatment units shows even distribution. As shown in Table 26 (p. 85), greedy matching did not yield a significant p-value. Whereas genetic matching produced a low p-value, indicating a significant difference between both groups. Fig. 39 depicts that genetic matching did not exclude high propensity scores thus producing an uneven match.
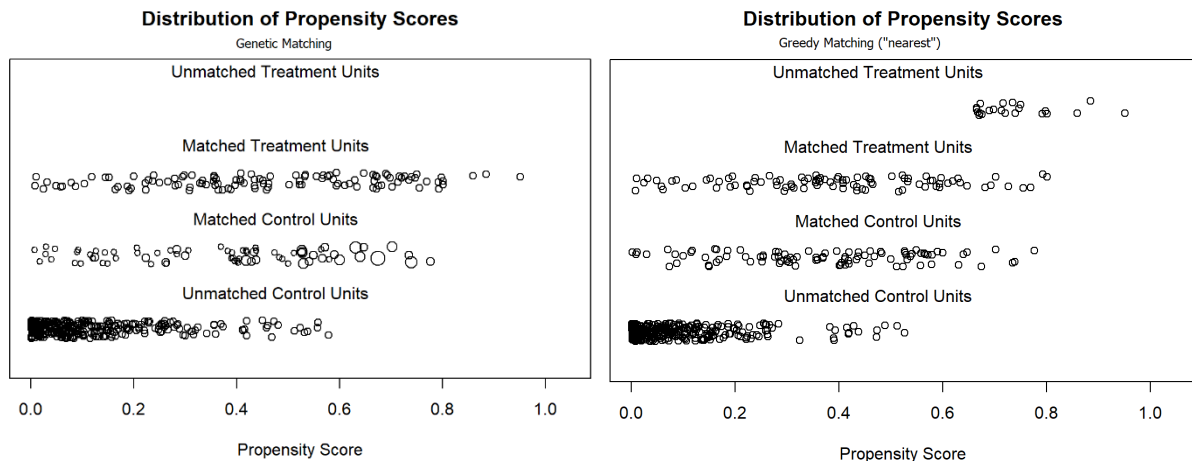
Fig. 39: Jitter Plots of Genetic and Greedy Matching (R Package *MatchIt*): Distribution of Propensity Scores

Fig. 40 demonstrates how genetic matching improves distribution of propensity scores for treated cases and controls. However, the frequency of propensity scores above 0.6 is higher in the treated than controls.
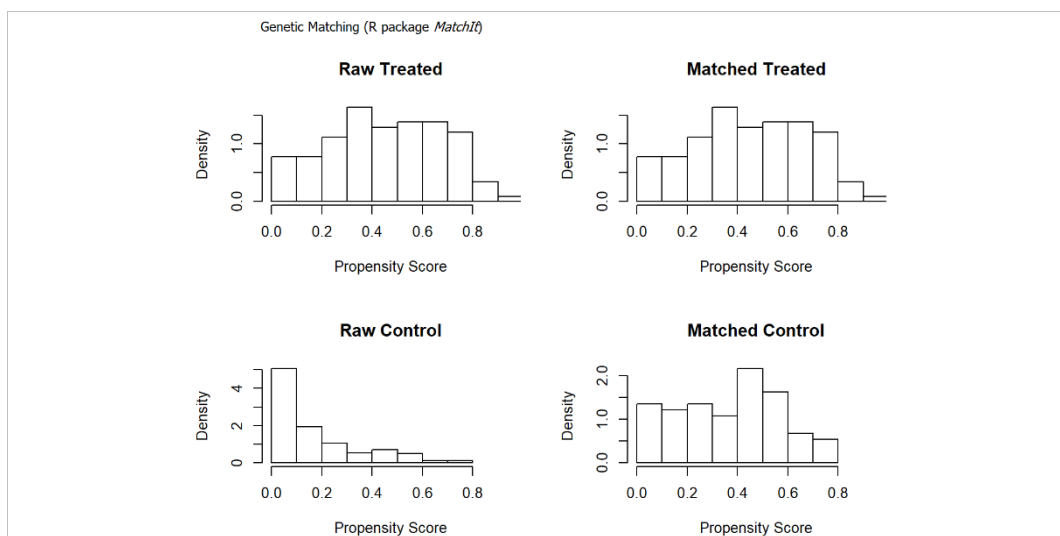
Fig. 40: Distribution of Propensity Scores before and after application of Genetic Matching (R Package *MatchIt*)

Finally, QQ plots, shown in Fig. 41, depict how matching improved the balance between treated and controls for many variables, however, perfect matching was not achieved.
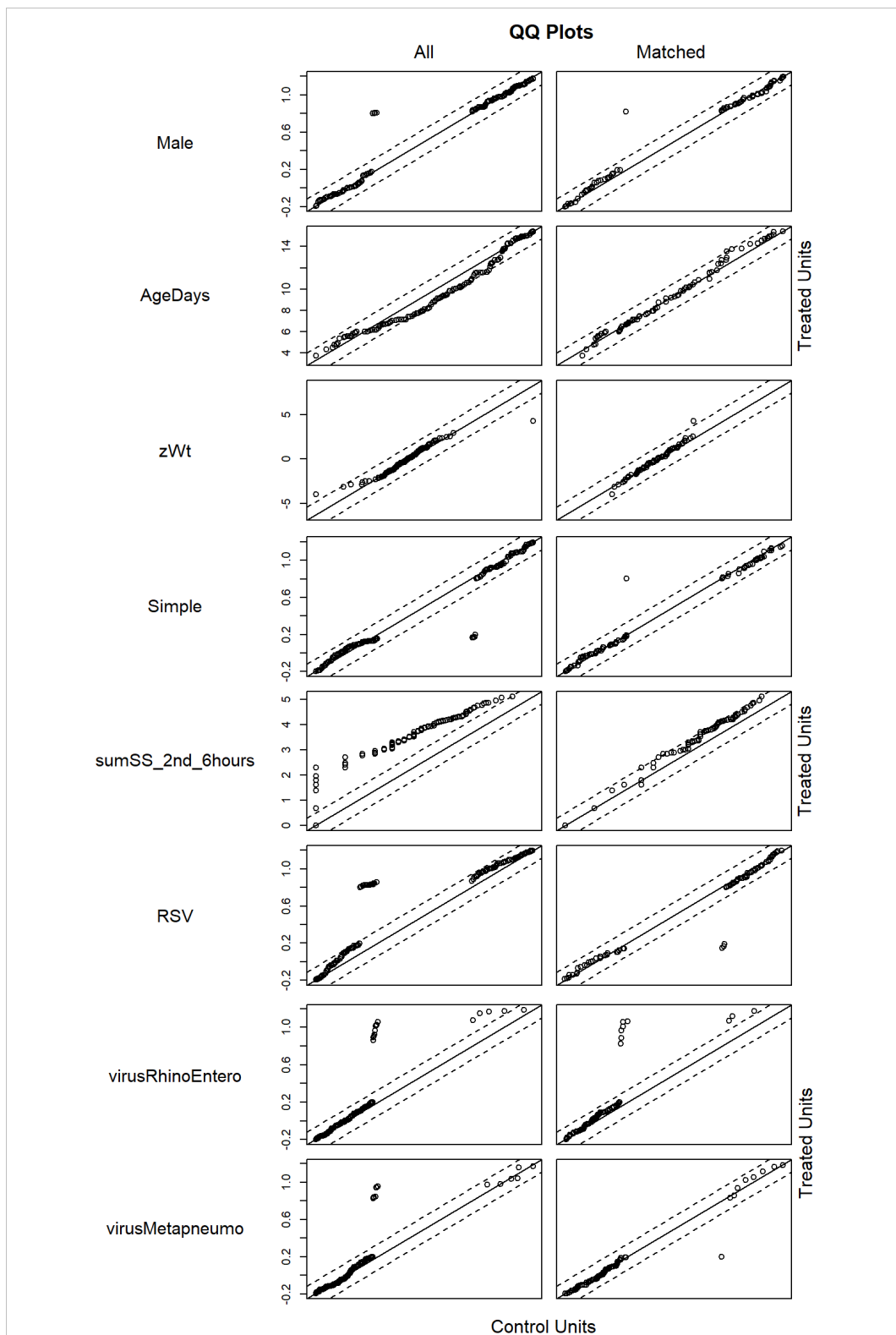
Fig. 41: QQ Plots Showing Results of Genetic Matching (R Package *MatchIt*)

7.5 Results of propensity score for different data subsets

Subgroup analysis was performed for age less than 100 days, respiratory severity score greater than 2.0 and respiratory severity score measured at the time of decision to treat. There was no significant effect of high flow therapy on the total length of stay found.

7.6 Discussion

As the main result of this chapter, propensity score matching demonstrated that patients treated with high flow therapy stayed between 10 and 16 hours longer in hospital than patients treated with standard care (=low flow oxygen). Most likely, this finding was the result of decision bias towards high flow therapy in patients with elevated respiratory severity scores thus indicating patients with severe disease. Because this study was based on observational, retrospectively collected data, the treatment decision was not randomized. Clinicians were more likely to initiate high flow therapy on patients with moderate to severe bronchiolitis than mild bronchiolitis. As a result, those patients were more likely to stay in hospital longer.

Several studies have investigated the influence of high flow therapy on hospital length of stay. High flow therapy was either compared with OxyMask delivery[110], nasal continuous positive airway pressure (nCPAP)[12,13,19,111,112], or standard low flow therapy.[16,18,51,113,114]

HFNC vs. OxyMask

Ergul and colleagues conducted a randomized controlled trial comparing high flow therapy and OxyMask (n=60, median age 10-11 months). They found that high flow therapy reduced the time in PICU. The hospital length of stay was also decreased, however, this was due to the fact that almost all patients who had moderate to severe bronchiolitis went through PICU.[110]

HFNC vs. nCPAP

In a small, retrospective study, Pedersen and colleague compared high flow therapy (n=27) with nCPAP (n=22). Both groups had similar hospital length of stay.[12] Milesi and colleagues randomly assigned 142 patients with acute viral bronchiolitis either to nCPAP or high flow therapy. They did not find a difference in hospital length of stay.[13] Three years earlier in a small retrospective study, Metge and colleagues came to the same conclusion (nCPAP n=19, HFNC n=15).[111] On the other hand, Guillot and colleagues identified a significant difference in hospital length of stay between exclusive high flow therapy (median 7 days; n=69) and nCPAP or biPAP (median 10 days, n=33).[19]

HFNC vs. Standard Therapy

With regards to comparison of high flow therapy with low flow oxygen therapy, most researchers did not find a significant difference in hospital length of stay, except Milani and colleagues, who found a three-day difference in the median hospital length of stay in favour of high flow therapy (median LOS: 9 vs. 6 days; n=36).[16,18,113,114]

In summary, three out of eleven studies found an advantage of high flow therapy over other non-invasive ventilation strategies when using length of stay as outcome variable. Three of the

RCTs that did not find a difference investigated large numbers of patients (n=142, n=1472, n=202).[13,16,18]

7.6.1 Limitations

The biggest limitation of the author's study was the single-centre, observational, and retrospective study design. To adjust for selection bias, feature selection and propensity score matching were used to find matching treatment and control groups. The level of matching achieved suboptimal levels by either omitting results from the treatment group or accepting persisting differences for some of the confounding variables. The difference in magnitude of the severity scores between the two groups was the most likely explanation for the current result.

7.6.2 Conclusion

This study's observational data were used to create a respiratory severity score which showed a significant positive correlation with hospital length of stay. The current analysis revealed a significant difference in severity scores between the high flow group and the standard group. Most likely, the reason why high flow therapy lead to prolonged length of stay was elevated levels of severity scores.

7.6.3 Future Research

These findings emphasize the need for studies that use a respiratory severity score for patient stratification. This score should be as objective as possible and fully validated. The decision to treat with low flow or high flow oxygen should incorporate the respiratory severity score.

The second, so far not fully answered, question is whether high flow therapy or nasal continuous positive airway pressure should be the first treatment in patients with moderate to severe bronchiolitis. Milesi and colleagues used a cross-over design allowing the switch between high flow therapy and nasal continuous positive airway pressure (nCPAP). Their management of secondary failure included BiPAP and high flow therapy at a higher flow level. This strategy kept the rate of conventional mechanical ventilation at a low level of 4.2% (nCPAP) and 6.9% (HFNC).[13]

8 Performance Analysis of Non-invasive Ventilation (NIV)

8.1 Introduction

So far, the results being presented dealt with variable selection and prediction of treatment and outcome. In addition, matching methods were used to find control and treatment groups that matched as good as possible to allow estimation of the effect of high flow therapy on length of stay.

In this chapter, the author analysed the reasons for success and failure of non-invasive ventilation in the standard group and high flow group. Failure was defined as either escalation from low flow to high flow nasal cannula or escalation from high flow nasal cannula to continuous mechanical ventilation (CMV).

The first step was to fit a generalised linear model (GLM) to calculate the significance level of important covariates in relation to failure of non-invasive ventilation. The list contained the following variables: age, gender, z score of body weight, gestational age, complex disease, RSV, and severity score. The severity score was extracted from three equally distributed time windows within each episode of non-invasive ventilation (Fig. 42).
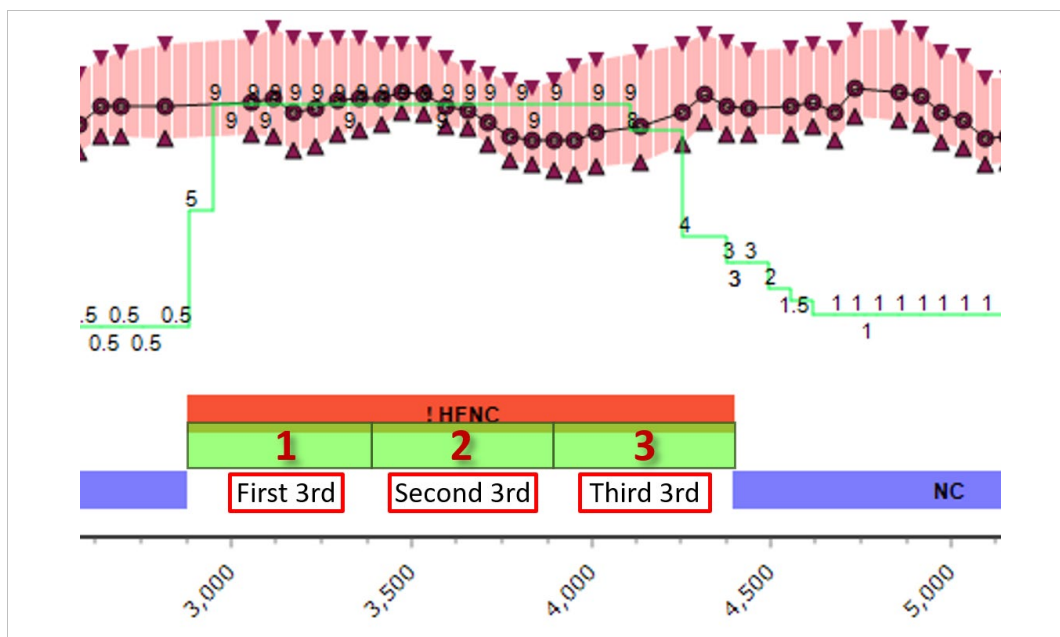
Fig. 42: NIV Relative Time Periods for Severity Score Analysis

The sum of the severity scores in one third of the successful group was compared with the sum of the severity scores in the corresponding third of the failure group. This comparative analysis was run on all three thirds in the standard group and high flow group, respectively. The clinical rationale for this three-part division was the assumption that the severity scores would increase at the end of a failing episode of non-invasive ventilation.

The second step was to apply the least square method for fitting a regression line; thereby calculating slope (M), correlation coefficient (R) and coefficient of determination (R²). This

technique was used to determine whether the slope of all severity scores measured during non-invasive ventilation was related to success or failure. Fig. 43 and Fig. 44 depict an example of each scenario (cases 708 and 916). The expected pattern for the success group was a negative slope indicating improving severity scores (Fig. 43).

Case 708 (SN=147-2016) LOS 5.52 days | M Age: 10 months 17 days
Admission: Z Score Weight 0.01, Body Weight 9.85 kg
Complex Case, HFNC

Best fit line Y=-0.0015 X + 11.6   M = -0.00149016613719298, X1 0 Y1 220 X2 860 Y2 310 | Slope -0.105
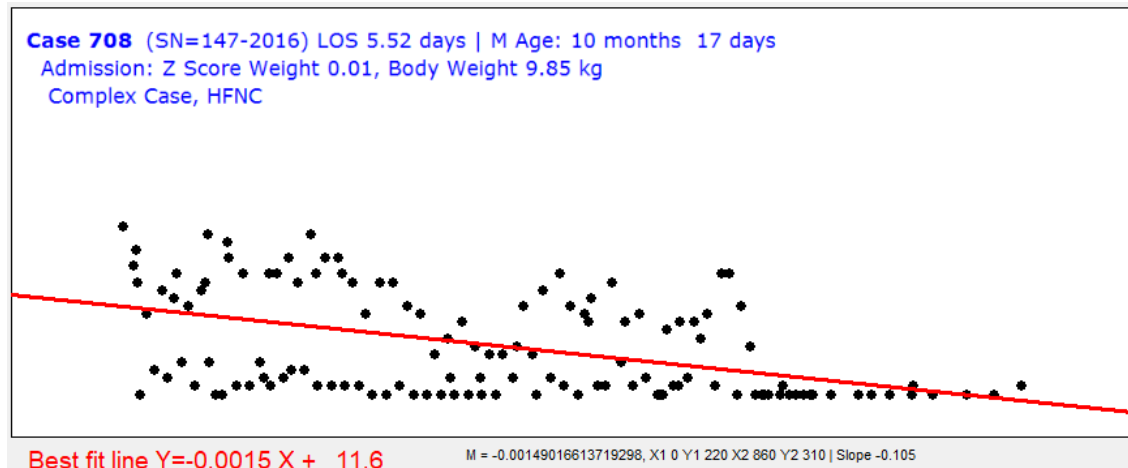
Fig. 43: Regression Line with Negative Slope Indicating Improving Severity Scores

The expected pattern for the failure group was a positive slope indicating worsening severity scores (Fig. 44).

Case 916 (SN=147-2016) LOS 9.17 days | M Age: 1 months 10 days
Admission: Z Score Weight 2.00, Body Weight 6.20 kg
Complex Case, HFNC escalated

Best fit line Y=0.0098 X + 4.1   M = 0.00975579419091959, X1 0 Y1 252 X2 860 Y2 146 | Slope 0.123
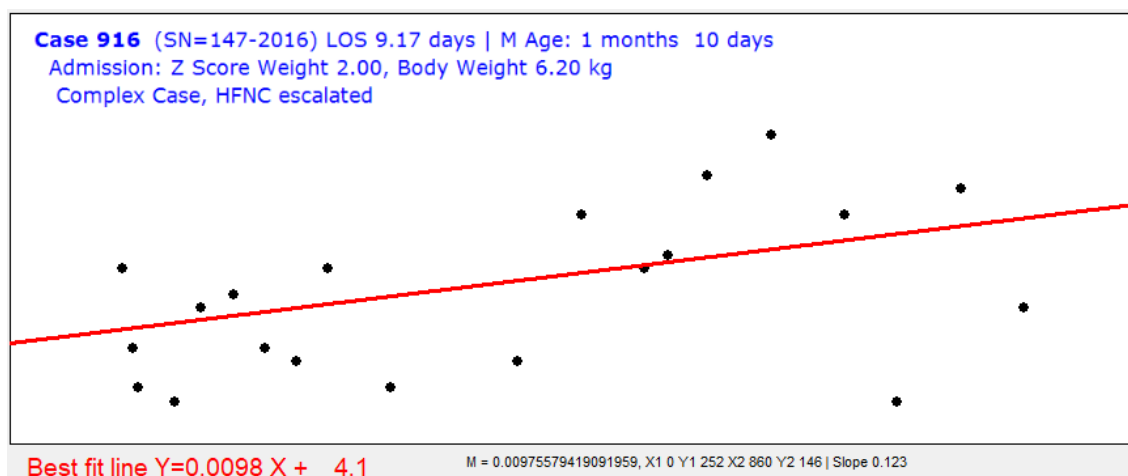
Fig. 44: Regression Line with Positive Slope Indicating Worsening Severity Scores

8.2   Factors that Predict Success and Failure of Non-Invasive Ventilation

8.2.1   Difference of Severity Scores in Relation to Treatment

A simple scatterplot of the respiratory severity score measured at the time of the treatment decision to commence either standard therapy or high flow therapy over the hospital length of stay divided in standard and high flow group, and successful and escalated cases is depicted in Fig. 45.
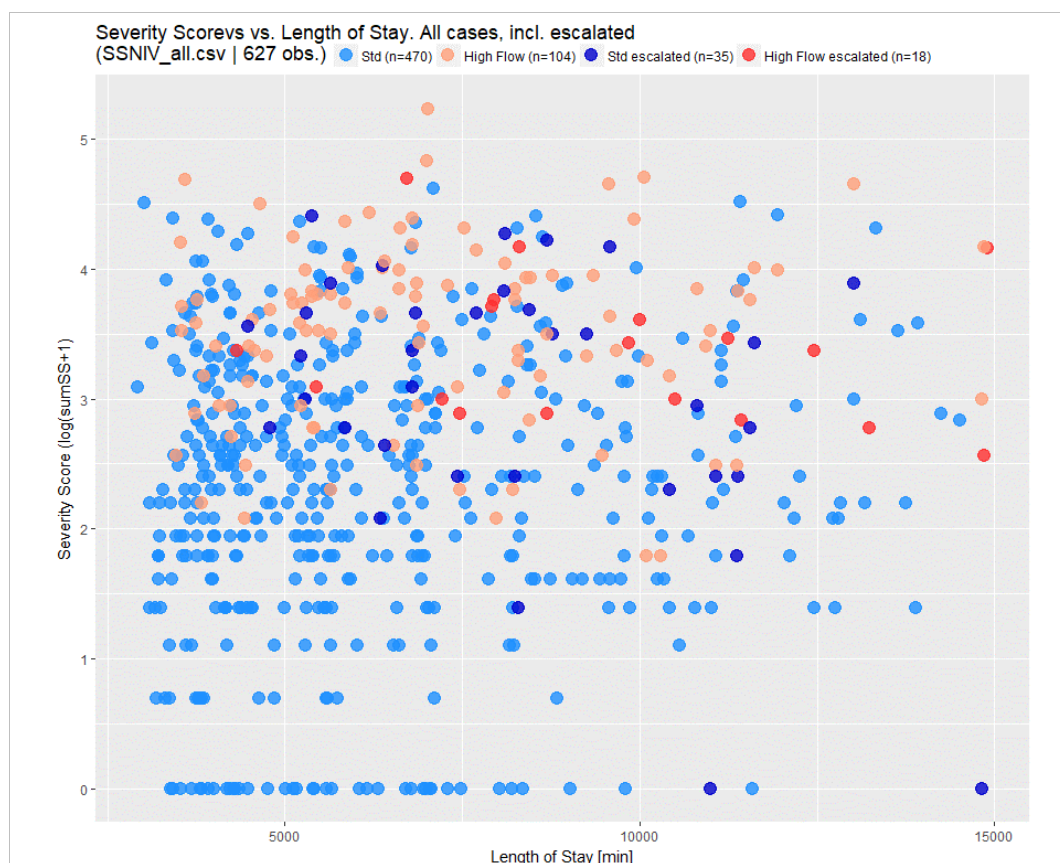
Fig. 45: Scatterplot Respiratory Severity Score Versus Hospital Length of Stay. All Therapies incl. Escalation.

The application of standard treatment (blue dots) occurred mostly at severity scores below three but also at higher levels, whereas the decision to treat with high flow nasal oxygen coincided, almost exclusively, at severity scores above two. The scatterplot shows that all escalated cases, no matter whether standard therapy (dark blue) or high flow therapy (dark red), started on an elevated severity score level. Apart from two outliers that had an artificially low severity score of zero (cases number 752 and 772), the escalated cases of standard therapy (dark blue dots) had started mostly at severity score levels of two and above. This finding was similar to the one of high flow. All cases of failed high flow therapy (red dots) started on high levels of severity scores (>2.5).

8.2.2   Difference in Severity Score within the High Flow Group

The following graph (Fig. 46) illustrates the relationship of the severity scores with the length of stay within the high flow group. The red dots depict the severity scores at the start of high flow therapy that was escalated to continuous mechanical ventilation (CMV). Escalation of high flow therapy occurred in 18 out of 122 cases (red dots, 14%).
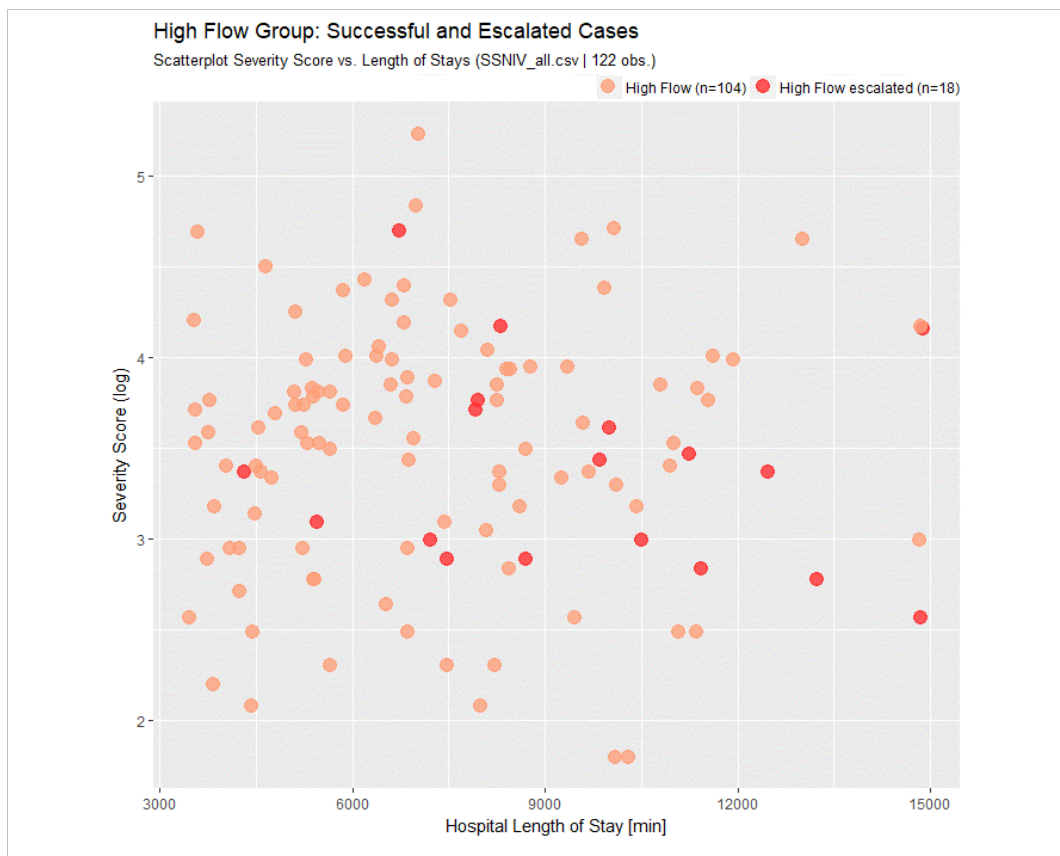
Fig. 46: High Flow Group: Successful and Escalated Cases. Severity Score versus Length of Stay.

There is no recognizable difference with regards to severity score between the two groups. When applying a generalized linear model with failed high flow therapy as outcome, only age had a significant negative correlation with escalation of high flow therapy (p=0.015), as shown in the next scatterplot (Fig. 47).
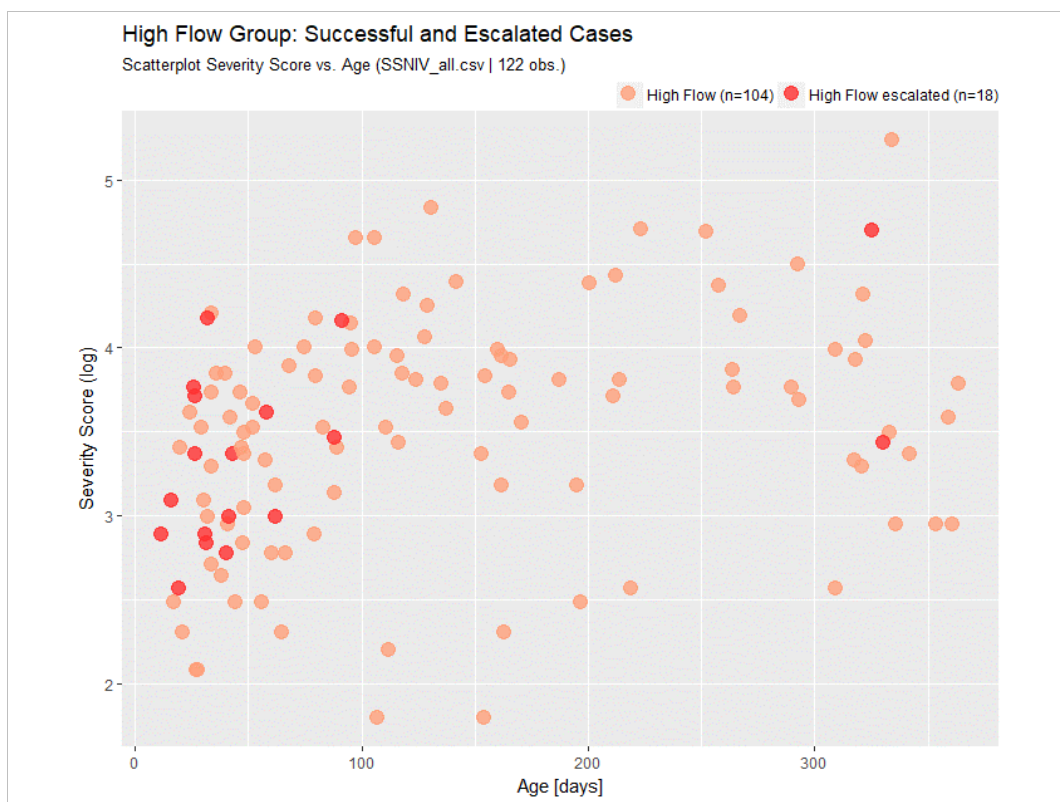
Fig. 47: High Flow Group: Successful and Escalated Cases. Severity Score versus Age.

In Fig. 47, the same severity scores were plotted but this time the age of the patient was used on the bottom axis. The distribution of the red dots demonstrated that 16 patients were younger than 100 days. Two patients were older than 300 days. Out of six covariates, only the odds ratio of age less than 100 days was significantly elevated: OR 12.78 at 97.5% CI 2.92 to 95 (Fig. 48).
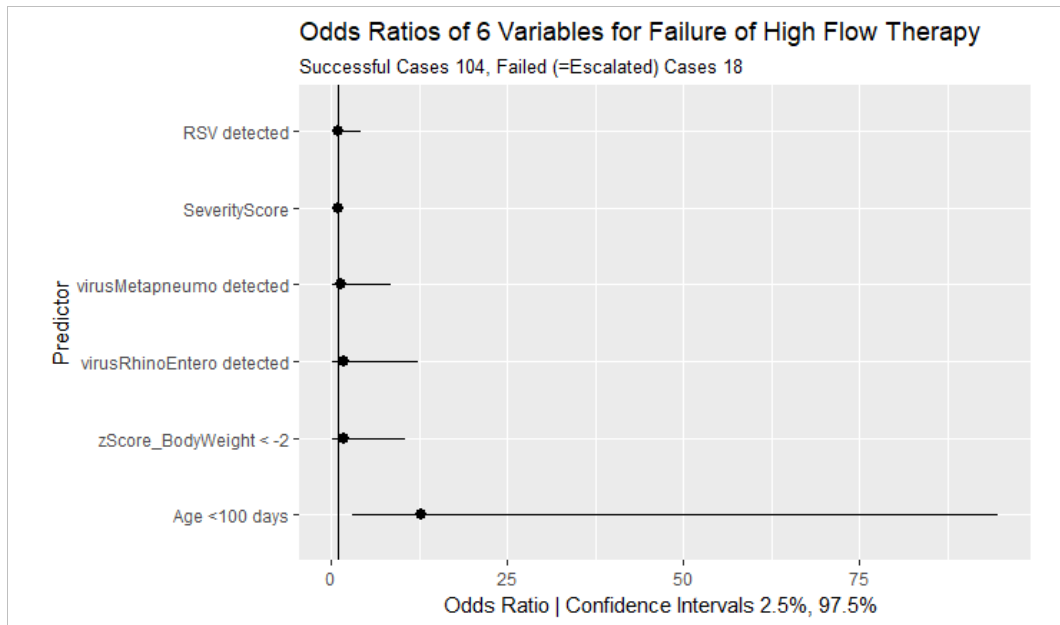
Fig. 48: High Flow Group: Odds Ratios of 6 Variables for Failure of High Flow Therapy

8.2.3   Difference in Severity Score within the Standard Group

Fig. 49 illustrates the relationship of the severity score with the length of stay within the standard group. The dark blue dots depict the severity scores at the start of low flow (NC) therapy that was escalated to high flow (HFNC). Escalation of low flow nasal oxygen occurred in 33 out of 503 cases (6.6%).
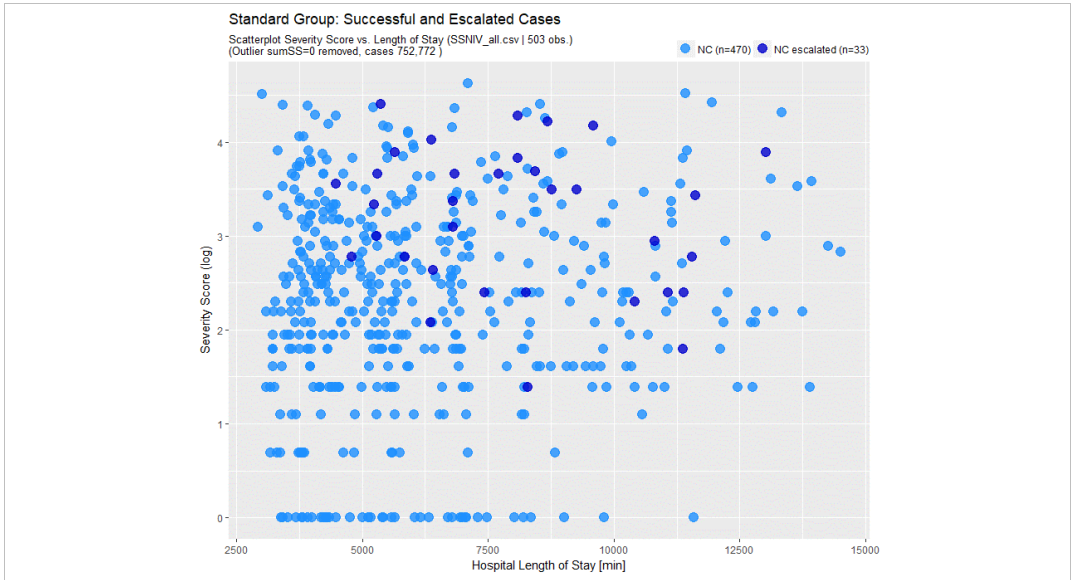
Fig. 49: Standard Group: Successful and Escalated Cases. Severity Score versus Length of Stay

Visual inspection of Fig. 50 suggested a positive correlation between respiratory severity score and failure of standard therapy. Application of a generalized linear model with failed standard therapy as outcome, revealed a positive correlation of severity score (p<0.001) and Metapneumo virus (p=0.0014). This time, age was not a significant (p=0.06) predictor for failure of standard therapy (Fig. 50).
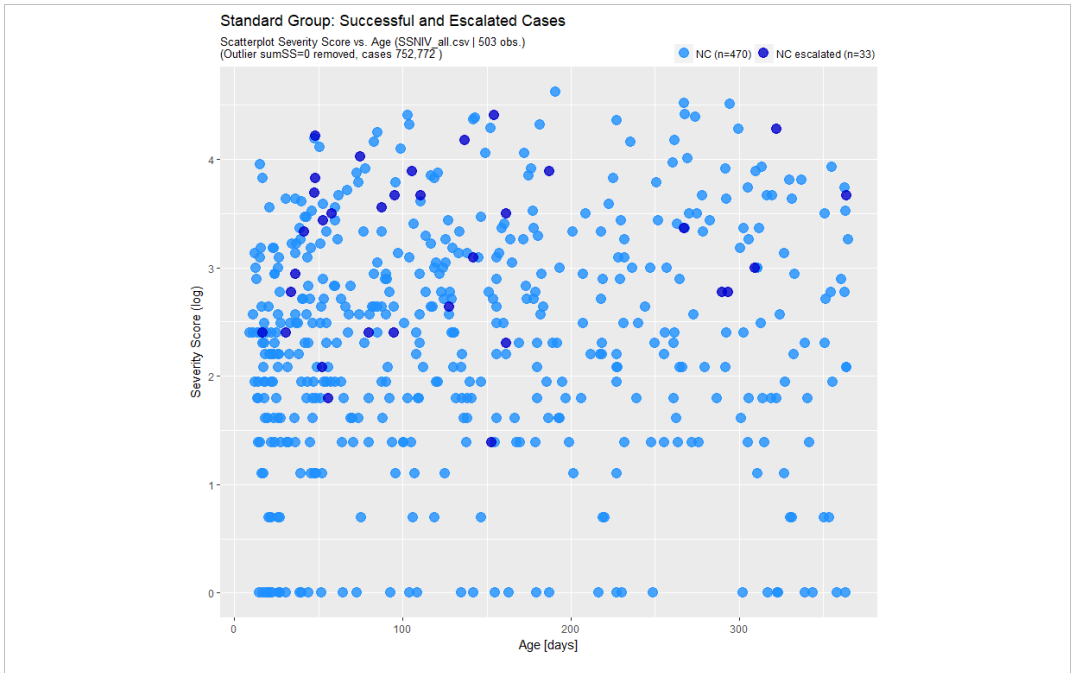
Fig. 50: Standard Group: Successful and Escalated Cases. Severity Score versus Age

To further investigate the influence of other covariates on failure of standard therapy, the odds ratios of severity score, age less than 100 days, z score of body weight, and viral cause (RSV, Metapneumo virus, Rhino-Entero virus) were calculated. As listed in Table 27, three predictors reached significance level: severity score greater than 2.0, age <100 days, and Metapneumo virus.

Table 27: Odds ratios to predict Failure of Standard Therapy

Predictor	Odds Ratio	CI_low (2.5 %)	CI_high (97.5 %)
SeverityScore >2.0	5.004	1.877	17.456
Age <100 days	2.546	1.109	6.201
zScore_BodyWeight < -2	2.728	0.797	8.202
RSV detected	1.364	0.607	3.196
virusRhinoEntero detected	1.818	0.271	7.262
virusMetapneumo detected	7.079	2.136	22.196

The graphical representation of the results is depicted in the following Fig. 51.
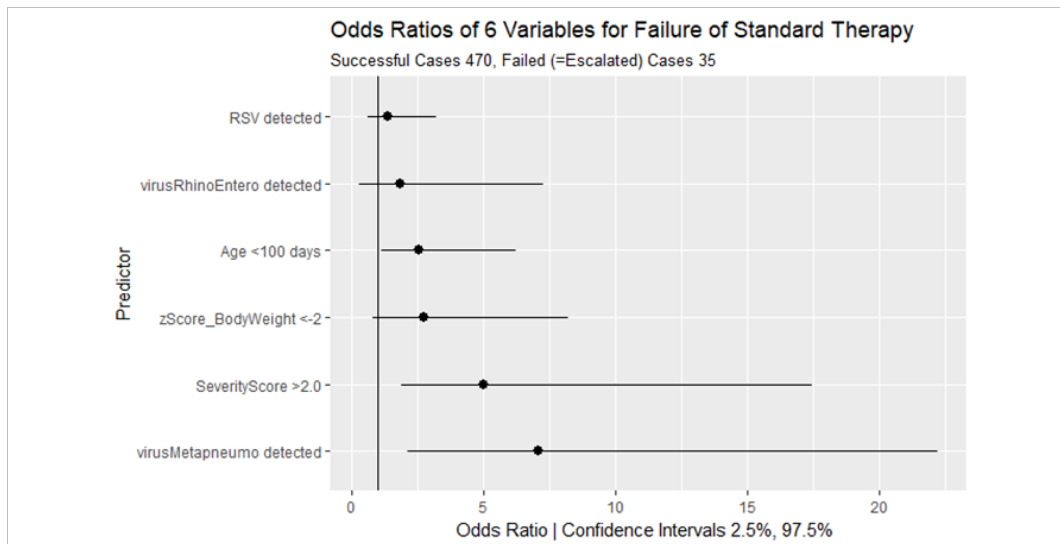
Fig. 51: Odds Ratios to predict Failure of Standard Therapy (Low Flow Nasal Oxygen)

8.2.4 Standard Group

The standard group had a failure rate of 7.4% (35 out of 470). Analysis of covariates in relation to success or failure did not reveal any significant influence of age, gender, z score of body weight, gestational age, complex disease, or RSV. These and the results for the three severity score divisions are shown in Fig. 52. When applying a generalized linear model, the results for the severity scores of the three thirds were statistically significant.

```
Escalated ~ Male + AgeDays + zWt + GA + Complex + RSV +
    ssNIV_1_First3rd + ssNIV_1_Second3rd + ssNIV_1_Third3rd

Call: glm(formula = mFormula, family = binomial, data = mdl)

Deviance Residuals:
    Min      1Q    Median      3Q       Max
-1.4305  -0.3603  -0.2318  -0.0668   3.5230

Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)      -0.238615   4.833662  -0.049  0.96063
Male             -0.442773   0.409742  -1.081  0.27987
AgeDays          -0.000567   0.002133  -0.266  0.79032
zWt               0.277891   0.165499   1.679  0.09313 .
GA               -0.007915   0.017169  -0.461  0.64480
Complex           0.537328   0.442233   1.215  0.22435
RSV               0.301691   0.433030   0.697  0.48599
ssNIV_1_First3rd  -0.028130   0.010195  -2.759  0.00579 **
ssNIV_1_Second3rd -0.056867   0.016594  -3.427  0.00061 ***
ssNIV_1_Third3rd   0.072622   0.012002   6.051 1.44e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 247.30  on 491  degrees of freedom
Residual deviance: 180.69  on 482  degrees of freedom
AIC: 200.69
Number of Fisher Scoring iterations: 8
```

Fig. 52: Failure in the Standard Group. Results of Logistic Regression Analysis for Nine Covariates

Sum Severity Score of the 1st, 2nd and 3rd Third

There was statistically significant, <u>negative</u> correlation between the severity scores and escalation of care for the 1st ($p<0.01$) and 2nd ($p<0.001$) third. There was significant <u>positive</u> correlation for the 3rd ($p<0.001$) third. However, visual analysis of the results, as depicted in Fig. 53, did not reveal any clear differences between successful and failed standard care. The number of zero outliers might have distorted the calculation. At this stage, it was decided not to investigate this further because of the small number of cases and missing values.
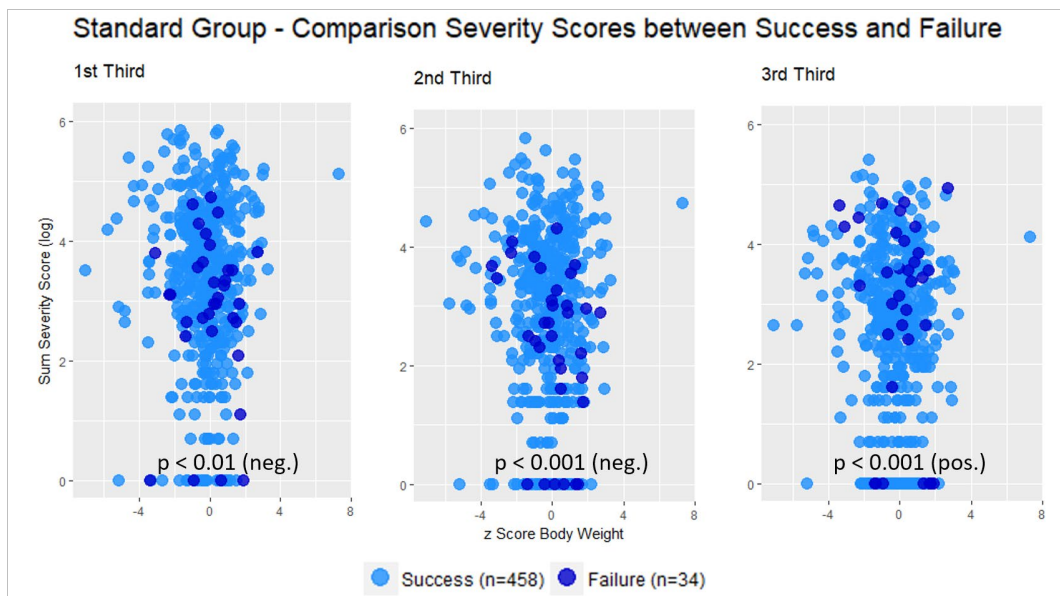
Fig. 53: Failure in the Standard Group. Severity Score of the 1st, 2nd, and 3rd Third

Slope of the Severity Score Regression Line

The second part of the performance analysis for the standard group looked at the significance of the slope of the regression line in conjunction with the other covariables mentioned above. As before, there was no significance of the six covariables but a significance level of $p<0.01$ for a positive slope of the regression line. Application of Extreme Gradient Boosting (XGBoost) revealed the following relative influence of all covariables (Fig. 54).
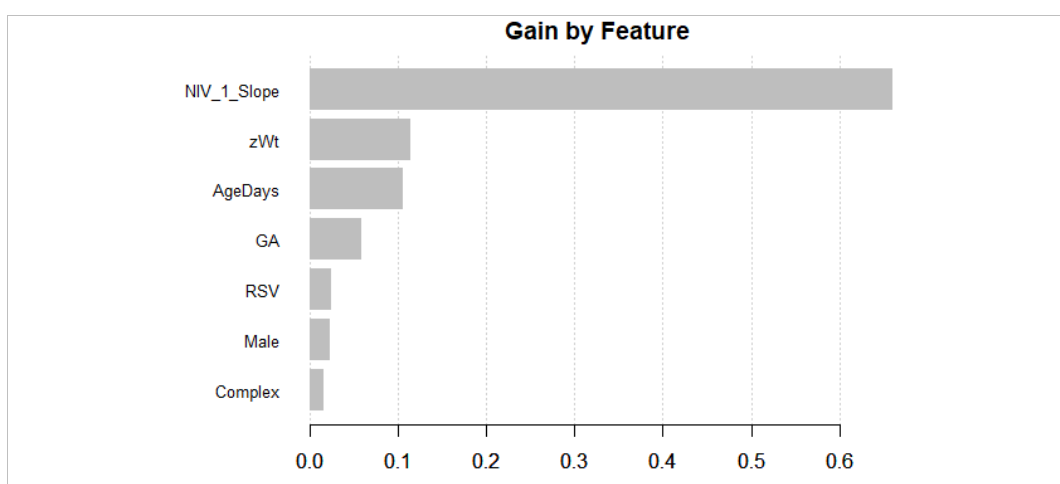
Fig. 54: Extreme Gradient Boosting for Covariables in the Standard Group in Relation to Failure

The scatterplot in Fig. 55 seemed to confirm this finding. The slope of the severity score was more positive in the failure group than success group. It appeared, that the z score of the body weight was lower in the failure group than success group.
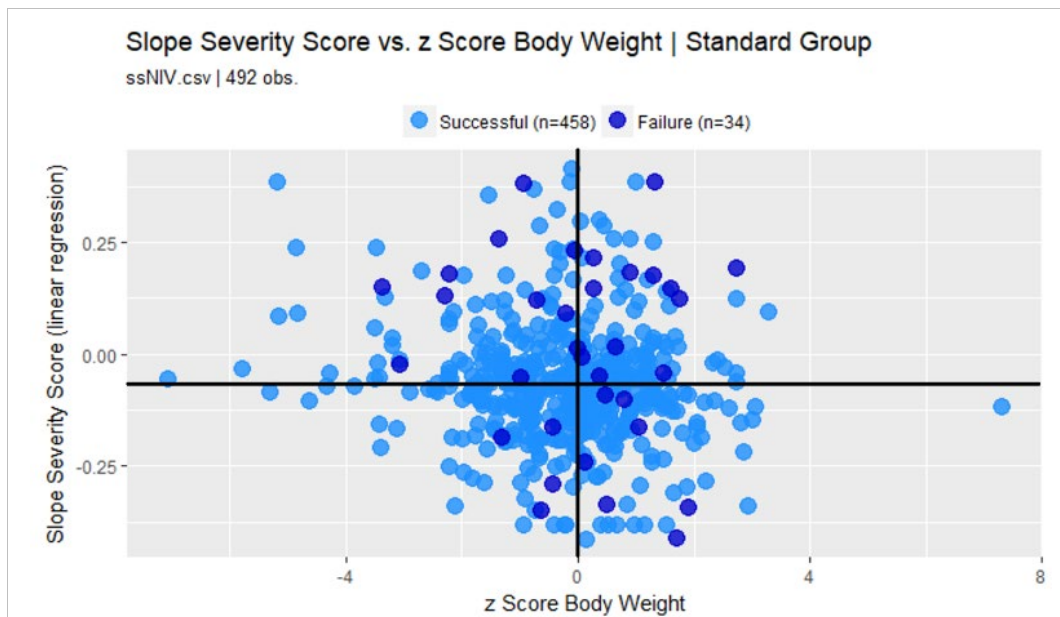
Fig. 55: Standard Group - Success/Failure. Scatterplot of Slope Severity Score vs. z Score Body Weight

8.2.5   High Flow Group

The high flow group had a failure rate of 14.8% (18 out of 122). Analysis of covariates in relation to success or failure did not reveal any significant influence of age, gender, z score of body weight, gestational age, complex disease, or RSV. The results for the three severity score divisions were similar except for the 1st third, which was not significant.

Sum Severity Score of the 1st, 2nd and 3rd Third

There was statistically significant, <u>negative</u> correlation for the 2nd third (p<0.05). There was significant <u>positive</u> correlation for the 3rd (p<0.01) third. Visual analysis of the results is depicted in Fig. 56. A small number of cases in the failure group, outliers and missing values make visual interpretation of the results difficult.
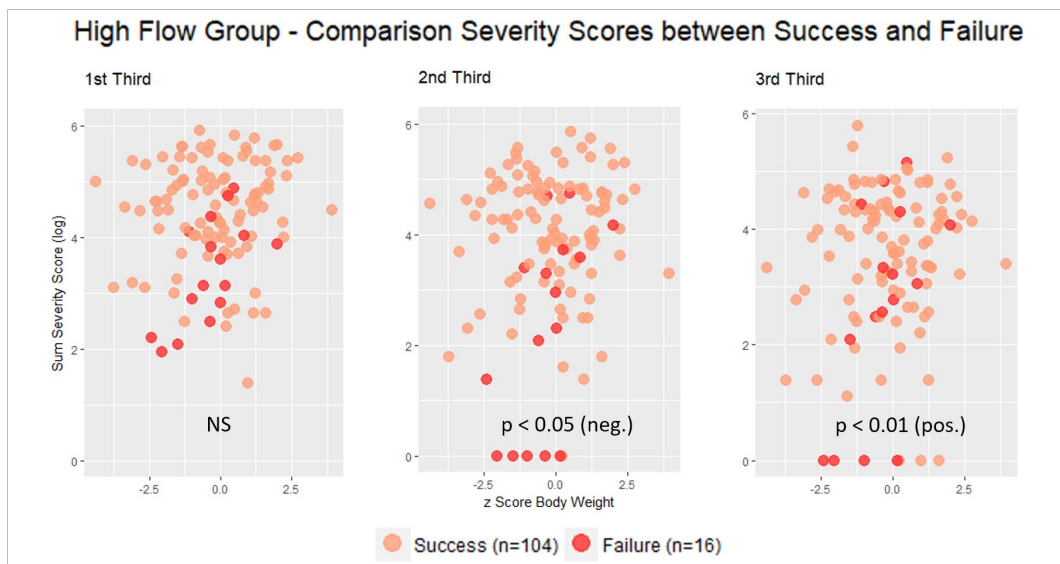
Fig. 56: Standard Group - Success/Failure. Scatterplot of Slope Severity Score vs. z Score Body Weight

Slope of the Severity Score Regression Line

The second part of the performance analysis for the high flow group looked at the significance of the slope of the regression line in conjunction with the other covariables mentioned above. As before, there was no significance of the six covariables but a significance level of $p<0.01$ for a positive slope of the regression line. Application of Extreme Gradient Boosting (XGBoost) revealed the following relative influence of all covariables (Fig. 57).
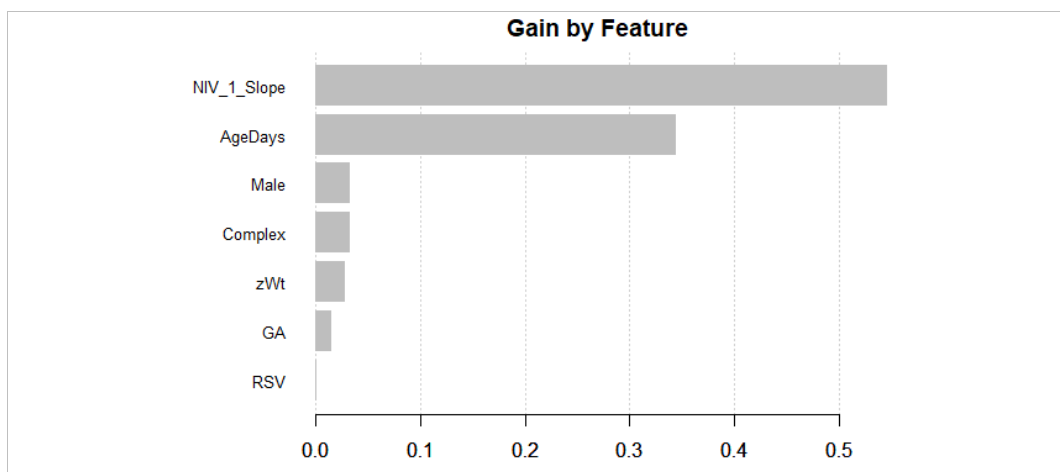
Fig. 57: Extreme Gradient Boosting for Covariables in the High Flow Group in Relation to Failure

The scatterplot in Fig. 58 seemed to confirm this finding. The slope of the severity score was more positive in the failure group than success group. Here again, the z score of the body weight was lower in the failure group than success group.
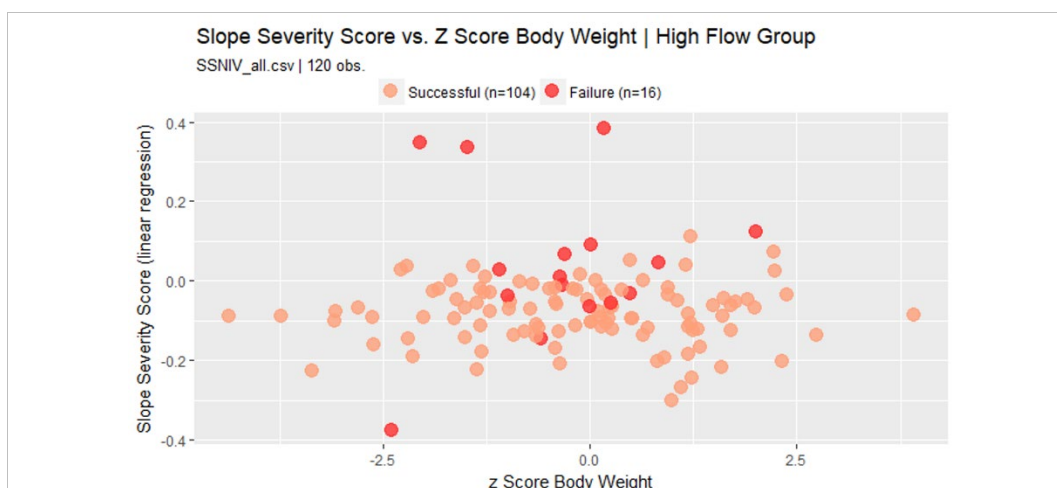
Fig. 58: High Flow Group - Success/Failure. Scatterplot of Slope Severity Score vs. z Score Body Weight

Flow Rate in the High Flow Group

To answer the question whether flow rate was correlated with failure, the *glm* function in R was used. It did not reveal any significant influence. Body weight or z score of the body weight were also not significant. Table 28 summarizes body weight and flow for the success and failure group. The median body weight was lower in the failure group than in the success group, however, the interquartile ranges were overlapping. The flow rate, adjusted for body weight, was similar in both groups. It was below the recommended rate of 2 L/Min/kg.

Table 28: Overview of Body Weight and Flow Rate in High Flow Group - Success/Failure

High Flow Nasal Cannula

Success Group, n=104

Body Weight [kg]					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.610	4.775	6.320	6.460	7.981	11.050

Average Flow [L/Min/kg]					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.09911	0.72797	1.07546	1.08041	1.41322	2.97672

Failure Group, n=18

Body Weight [kg]					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.440	3.651	4.049	4.809	5.357	9.900

Average Flow [L/Min/kg]					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3439	0.9366	1.1790	1.2433	1.5765	2.5046

8.3 Discussion

The main result of this chapter is that respiratory severity scores strongly correlate with success or failure of non-invasive ventilation. There was statistically significant, negative correlation between the respiratory severity scores and escalation of care for the first and second third in the standard group, and only for the second third in the high flow group. Both treatment groups had a significant positive correlation for the 3rd third for treatment escalation. In addition,

the slope of the regression line of the severity score was more positive in the failure group than success group. This finding was statistically significant for the standard and the high flow group.

Another result of this chapter is that age was a highly significant predictor of failed high flow therapy. The odds ratio of age younger than 100 days was 12.8.

Looking at failure of standard therapy, a high respiratory severity score, obtained shortly before initiating standard therapy, was significantly correlated with failure of standard therapy. The scatterplot on page 95 (Fig. 49) suggested that a severity score of more than two significantly increased the chance of treatment escalation. The odds ratio was 5.0 (97.5% CI, 1.88 – 17.46). Age younger than 100 days and Metapneumo virus were also positively related.

Several studies have investigated predictors of failure of high flow therapy. Abboud and colleagues found in a retrospective chart review that 21 (18.6%) out of 113 bronchiolitis patients admitted to paediatric intensive care required continuous mechanical ventilation. The analysis revealed predictors of failure as follows: lower respiratory rates, higher pCO_2 levels, higher PRISM scores (Pediatric Risk of Mortality III) and no decrease in respiratory rate following initiation of high flow therapy. Age and history of prematurity did not correlate with failure. Due to the small number of cases and the limited number of variables investigated, generalizability of the results was limited. There was no statistical model and performance parameters (e.g. ROC AUC, specificity, sensitivity) provided that evaluated the overall predictive value of all covariates.[43]

Kelly and colleagues conducted a retrospective cohort review of 498 children younger than two years who presented to the emergency department with respiratory distress and received high flow therapy within 24 hours of initial triage. The causes for respiratory distress was bronchiolitis, pneumonia and asthma. The variables that predicted failure of high flow therapy were as follows: admission respiratory rate greater than 90th percentile, initial venous pCO_2 greater than 50 mmHg, and initial venous pH less than 7.3. Out of 216 patients with viral bronchiolitis, 15 (7%) required continuous mechanical ventilation. The main result of this study was the high success rate of high flow therapy. Because of missing data (38% of the cases had to be excluded for final analysis) and small number of variables, the generalizability of the results was limited.[115]

Betters and colleagues identified in a retrospective chart review of 192 patients with primary respiratory disease significant risk factors associated with failure of high flow therapy as follows: history of cardiac disease, history of prior intubation, a non-bronchiolitis diagnosis, and elevated FiO_2 needs (>0.5). Out of 142 patients with acute viral bronchiolitis five (3.5%) required continuous mechanical ventilation. History of prematurity and admission weight did not associate with higher failure rates. The severity of disease was not evaluated.[116]

A small study from Melbourne, Australia, examined 71 patients who received high flow therapy in the emergency department. Out of 49 patients with viral bronchiolitis 21 (43%) eventually required continuous mechanical ventilation. Even though vital signs such as heart rate, respiratory rate, body temperature, and oxygen saturation were recorded, no attempt was made to identify predictors of failure.

More recently, Kepreotes and colleagues randomly assigned 202 bronchiolitis patients to either standard therapy or high flow therapy. The failure rate for standard therapy was 33% (n=33) and 14% (n=14) for high flow therapy. High flow therapy was used in 61% (n=20) of the 33 cases who experienced treatment failure on standard therapy. The time to treatment failure was longer in the high flow group than standard group. There was no difference between the groups in relation to time on oxygen. The authors concluded that high flow therapy might have a role as first-line rescue therapy for patients treated with standard low flow. Kepreotes et al. used a flow rate of 1L/Kg/min for high flow therapy which is below the now recommended 2 L/kg/min.[18,117]

Heikkilä and colleagues retrospectively investigated bronchiolitis patients less than one year old treated at six Finnish hospitals. Out of 88 patients, 12 (13.6%) patients were either escalated to continuous mechanical ventilation (n=2) or nasal CPAP (n=10). Birthweight and gestational age were significantly lower in failed cases than successful cases. As a possible predictor of high flow failure, lack of improvement of heart rate and oxygen saturation during the first 60 minutes of high flow treatment were mentioned. Because improvement of respiratory rate took much longer, it was not considered an early predictor of failure.[118]

Guillot and colleagues conducted an observational prospective study in a paediatric intensive care unit examining 102 bronchiolitis patients. They compared two study periods, before and after the introduction of high flow therapy as the first line treatment. The failure rate in the second period was 38% (21 out 55 patients). The only independent factor for failure was higher pCO_2.[19]

In this study, the overall failure rate of standard or high flow therapy was below the rate reported in the literature. The treatment failure occurred more frequently in the high flow group (18 out of 122, 14.8%) than in the standard group (35 out of 470; 7.4%). Compared with the PARIS trial[16], the percentages seemed to be reversed: 7.4% in the standard group versus 14.8% in the high flow group which stands in contrast to the main result of the RCT from Australia and New Zealand. Most likely, the observational nature of the current study and complexity of diagnostic codes accounted for the main difference between the author's patient cohort and the one from Australia/New Zealand. The prospective design of the PARIS trial allowed correct identification of the diagnosis at the time of enrolment. In Los Angeles, the only source of diagnostic information was ICD-9 and ICD-10 codes to identify patients with the principal diagnosis of acute viral bronchiolitis. As per study design, other respiratory and non-respiratory conditions were permitted.

Therefore, for some parts of the analysis, the distinction was made between simple and complex cases. The attribute *simple* was applied to cases with bronchiolitis as the only diagnosis.

Table 6 (Patient Characteristics, p. 36) demonstrated that in the group of simple cases the relative numbers of cases and the numbers of treatment escalations were nearly equal between standard and high flow group: 3.4% in the standard group (n=16) versus 4.1% in the high flow group (n=5). In the high flow group, escalation of care occurred mainly in complex cases. The pronounced difference between the groups in relation to "Other diseases of the respiratory system" (ICD Codes J96 - J999) could be the explanation for the higher escalation rate in the high flow group.

There was potentially another reason why this retrospective, observational study found a higher escalation rate in the standard group than in the PARIS trial. Fig. 45 on page 92 depicted the severity scores at the beginning of the first relevant non-invasive ventilation episode for each case. The difference in severity score between standard and high flow group was highly significant, most likely reflecting the clinician's decision to either commence standard therapy or high flow therapy. It is reasonable to assume that high flow therapy would have been applied only to very sick patients. On the other hand, Fig. 45 showed that 35 patients received standard therapy, despite a high severity score. One could speculate, that this was perhaps the reason why their treatment failed; their score was too high in the first place. There was only one patient who was escalated to invasive ventilation, the remaining 34 were escalated to high flow therapy.

9 General Discussion

This retrospective, observational study used electronic health records of a large tertiary paediatric hospital in the U.S. to compare how high flow oxygen application via nasal cannula with standard low flow oxygen application in the management of acute viral bronchiolitis in infants. Propensity score matching demonstrated that patients treated with high flow therapy stayed between 10 and 16 hours longer in hospital than patients treated with standard care, i.e. low flow oxygen. Most likely, this finding was the result of decision bias towards high flow therapy in patients with elevated respiratory severity scores thus indicating patients with severe disease. Because this study was based on observational, retrospectively collected data, the treatment decision was not randomized. Clinicians were more likely to initiate high flow therapy on patients with moderate to severe bronchiolitis than mild bronchiolitis. As a result, those patients were more likely to stay in hospital longer.

Another result of this study was identification of main predictors of hospital length of stay. The highest-ranking individual parameters of respiratory distress to predict hospital length of stay were as follows: capillary pO_2, signs of retractions and respiratory effort, breath sounds, heart rate, body temperature, respiratory rate, and pulse-oximetry. These and other pertinent parameters were subsequently used to calculate a data-driven respiratory severity score. A computer algorithm used pre-defined rules to convert the value of each individual item into a weighted score. The sum of the individual scores were then used to describe severity of illness at any given point in time.

The newly created respiratory severity score was significantly correlated with total length of stay, high flow therapy, and success rate of standard and high flow therapy. The severity score predicted total length of stay when measured at the time of the treatment decision to apply either standard care or high flow therapy ($p<0.001$, odds ratio of 1.215 to predict LOS>5 days) or calculated from the second six-hour period following the first recorded event ($p<0.01$). The severity score predicted high flow therapy with a high area under the receiver operating characteristic curve (ROC AUC) of 0.84 when applying a Naïve Bayes model. When using the severity score at time of treatment decision in conjunction with six other covariates, a generalized linear model (GLM) was the best fitting model to predict high flow therapy (ROC AU 0.83, sensitivity 0.94, specificity 0.33). The odds ratio calculated 2.8 (97.5% CI, 2.23 to 3.52).

This study used methods of machine learning to identify main determinants of treatment success and failure. The severity score demonstrated a significant correlation with success and failure rate of standard and high flow therapy. The episodes of non-invasive ventilation analysed in this study were divided into three thirds in order to investigate cumulative changes of respiratory severity scores. The Severity score of the first and second third in the standard group and of the second third in the high flow group was <u>negatively correlated</u> with failure of treatment, whereas

high severity scores of the 3rd third in the high flow and standard group predicted the need to escalate the treatment. In addition, the slope of the regression line of the severity score during non-invasive ventilation was significantly more positive in the failure group than success group.

This study confirmed the known influence of age and physical development on length of stay and high flow therapy. Corrected age at the time of admission was an important predictor for total hospital length of stay. Its odds ratio showed a negative relationship of 0.922 to predict prolonged hospital length of stay, i.e. greater than five days. Age younger than 100 days showed an odds ratio of 2.97 (97.5% CI; 1.737 to 5.19) to predict high flow therapy.

Z score of the body weight was another important predictor for total hospital length of stay. Its odds ratio showed a negative relationship of 0.875 (97.5% CI; 0.816 to 0.936) to predict prolonged length of stay of more than five days. Z score of the body weight of less than -2 showed an odds ratio of 2.57 (97.5% CI; 1.175 to 5.47) to predict high flow therapy.

Cases without comorbidities, i.e. acute viral bronchiolitis as the only diagnosis, had an odds ratio of 0.403 (97.5% CI; 0.265 to 0.606) to predict prolonged LOS (>5 days).

Finally, the influence of various viral causes on prolonged LOS (>5 days) was calculated. Metapneumo virus had a significant odds ratio of 2.322 (97.5% CI; 1.065 to 4.995) whereas the influence of RSV did not show significant correlation. RSV accounted for 50% and Metapneumo virus only 6.3% (total 574 cases). There were increased odds ratios of RSV, RhinoEntero virus and Metapneumo virus to predict high flow therapy.

Discussion of Literature

Comparative effectiveness research (CER) in paediatrics is not widely practiced yet. In the following, two recent articles are mentioned which published new findings based on the analysis of observational data. A retrospective comparative study of antibiotic therapy delivered via the peripherally inserted central catheter and the oral route after hospital discharge in children with acute osteomyelitis demonstrated no advantage of the invasive route.[119] Another retrospective, observational study using propensity score matching compared the outcomes of high frequency oscillatory ventilation (HFOV) with those of conventional mechanical ventilation (CMV) in children with acute respiratory failure. The authors of the study concluded that the use of HFOV was associated with worse outcome; however, several commentators pointed out that important predictors of HFOV had not been part of the dataset therefore distorting the result.[120,121] The use of the correct methodology, in this case propensity score matching and sensitivity analysis, cannot overcome the problem of incomplete or inaccurate data. These articles serve as an example on how large scale EHR analysis has the potential to complement randomized controlled trials (RCTs).[122] This study was not able to present a result that was on the same level of evidence as a randomized controlled trial because selection bias made it impossible for propensity matching to find well

balanced treatment and control groups. The high flow group was significantly different from the standard group.

One reason that randomized controlled trials provide the highest level of evidence is that they avoid selection bias with regards to treatment assignment. In the case of acute viral bronchiolitis, the level of severity can vary substantially. If randomization is based only on the presence of the disease and not severity, and the number of severe cases is less frequent than mild to moderate cases, then interpretation of the results becomes difficult. Korppi asked in his editorial, after reviewing the studies by Kepreotes et al.[18] and Franklin et al.[16], "Should we routinely use high-flow oxygen therapy instead of low-flow oxygen administration for all bronchiolitis patients who need oxygen?" His answer was "No", because two-thirds of the bronchiolitis patients who needed oxygen supplementation were successfully treated with standard low-flow therapy. In addition, high flow therapy worked well as a rescue therapy.[20] The use of a validated and standardized severity score would assist with the discrimination between the therapeutic need for high flow or low flow.

9.1 Strengths

The main strength of this study is its size. A large dataset, extracted from a production database that represented real-world data, was used to create a data-driven respiratory severity score. A computer algorithm applied a pre-defined ruleset to structure and analyse relevant data in a uniform way. The respiratory severity score combined objective data, e.g. age, z score of body weight, heart rate, respiratory rate, oxygen saturation, capillary pO_2, body temperature, with subjective assessments, e.g. signs of retractions and respiratory effort, breath sounds. Full validation could not be achieved. However, the newly created severity score demonstrated promising characteristics when used in a fully computerised healthcare setting. Face and content validity met the Bekhof Criteria (Appendix C). In addition, construct and prediction validity were successfully evaluated by applying basic statistics and machine learning tools.

9.2 Limitations

The main limitation of this study was its partial reliance on manually entered data. Manual entry of signs of dyspnoea was performed by a variety of clinical staff. This included respiratory physicians, emergency and intensive care physicians, as well as nurses and respiratory technicians. Measures of inter- and intra-observer reliability and internal consistency could therefore not be evaluated. The current dataset was not precise enough for formally assessing responsiveness. In addition, the categorical fields contained manually entered text which was more difficult to analyse because methods of natural language processing (NLP) were not applied. Some of the text items were redundant and sometimes not entered consistently. The existing data fields were not designed

to be scientifically or logically correct. Many of the fields were designed to comply with legal requirements for documentation or were designed for certain areas such as emergency department or intensive care. Therefore, it was not attempted to determine optimal cut-off values for the severity score.

Another limitation was the usage of ICD codes (International Code of Diseases versions 9 and 10). The search algorithm identified all patients with acute viral bronchiolitis, however, it was not always evident whether acute viral bronchiolitis was the principal diagnosis. As part of the cleaning process, many cases had to be manually excluded. In addition, the researchers of this study had to decide whether acute viral bronchiolitis was the main determinant for hospital length of stay. For example, patients who required heart surgery because of congenital heart disease or patients who suffered from severe respiratory problems were excluded. This decision was not always easy to make because of incomplete or missing data. It was pragmatically assumed that every hospital length of stay greater than twenty days could not be contributed to acute viral bronchiolitis. Analysis of outliers for hospital length of stay lead to further restriction of the dataset to hospital length of stay less than 15,000 minutes, i.e. less than 10,42 days.

Another limitation was the lack or paucity of data regarding the influence of known risk factors, e.g. family smoking habits, family history of asthma, lack of breast feeding.

9.3   Future Research

Full Validation of respiratory severity score

The literature review (chapter 2, p. 18) highlighted the importance of a validated severity score for clinical research. The treatment effect of an intervention cannot be evaluated without assessing the degree of severity at the start of the intervention. The literature on acute viral bronchiolitis often points out that the severity of illness was mild to moderate without defining how it was measured, or non-validated measurement tools are applied. The fact that a measurement tool is not validated does not mean that it is wrong, but testing all three domains of quality criteria (validity, reliability, utility) ensures reproducibility and generalizability of the results.[34]

The design of a prospective, observational study should include blinded testing of inter-observer and intra-observer variability. The use of a stethoscope would be either reduced to a minimum or sufficient training prior to study start would have to be applied. The text items should be designed without overlap of meanings. Further research should determine the minimum requirement of items.

Finally, the general applicability of the respiratory severity score for other disease entities such as asthma or pneumonia, or age groups other than infants should be tested.[123]

Utilisation of respiratory severity score as a decision tool

A fully validated respiratory severity score could be used to assist with the treatment decision whether a patient actually requires non-invasive ventilation. Further research needs to evaluate at what severity level which mode of non-invasive ventilation should be applied, either standard nasal oxygen, heated and humidified high flow nasal oxygen, nasal continuous positive airway pressure, or conventional mechanical ventilation. In addition, a fully validated respiratory severity score can be used to calculate optimal cut-off values regarding these treatment decisions and prediction of outcome.

This and other studies showed that high flow therapy represented an excellent rescue therapy for standard nasal cannula (34 out 35, 97%). So far, the question whether high flow therapy or nasal continuous positive airway pressure should be used as first line treatment has not been answered.[124] In addition, Wang and colleagues concluded that a certain degree of experience is needed to decide which patient would benefit from high flow therapy or other modes of non-invasive ventilation. The respiratory severity score described in this study might assist the clinician to answer this question.

Optimizing electronic health record data entry and analysis

Currently, most electronic health record systems are optimized for billing and documentation but not for statistical analysis and scientific research. At present, electronic health record systems have limited ability to provide clinical decision support that is based on real-time data. Part of the challenge is that the data cleaning process is complicated and time-consuming. Further research is needed to find out which algorithms and methods of data entry optimally assist with the goal of real-time data processing and clinical decision making. "Intelligent" EHR software needs to provide helpful dialogs for meaningful and fully validated data input. To avoid errors of data entry, physiologic limits should be considered. For example, testing z scores at the time of data entry might warn the user of highly unlikely values.

Most recently, Rubin and colleagues described data-driven, machine learning algorithms to predict patient transfers to paediatric intensive care. They found improved performance parameters (accuracy, sensitivity, specificity, AUROC) when comparing the machine learning tool with modified paediatric early warning score (PEWS).[125] Whereas Ross and colleagues did not find improved performance parameters when integrating data-driven heart rate and respiratory rate into a validated paediatric early warning system without using machine learning tools.[126] A multitude of studies is required to determine which data is required to detect early deterioration of the paediatric patient.

Luo and colleagues conducted a systematic literature review on predictive modeling for bronchiolitis. Out of 2312 references 168 were determined to be relevant and were discussed in

the review. He concluded that many existing models had inadequate accuracy, did not cover some parts of their full scope, or, for some issues related to bronchiolitis, no predictive model had been built. He pointed out that no attempts have been made so far to translate a predictive model for bronchiolitis into routine clinical use and then re-evaluate patient outcome.[98] More research is required in this area.

9.4 Conclusion

To the author's best knowledge, this is the first study to describe a data-driven severity score for acute viral bronchiolitis in infants. It incorporated eleven numerical and eleven categorical values which were converted into individual, weighted scores. The sum of the individual scores generated the respiratory severity score. Paediatric early warning systems have been implemented in many paediatric healthcare facilities, however, there is lack of evidence of their effectiveness. This study demonstrates how EHR-driven scoring of the patient's condition might assist with treatment decisions and predictions.[126-128]

Respiratory severity scores, obtained from various time points of the hospital stay, and other covariates (risk factors) were used to fit machine learning models that predicted hospital length of stay, prolonged length of stay (>5 days), the need for high flow therapy, and failure of standard therapy or high flow therapy.

Due to decision bias, propensity score matching could not demonstrate a treatment effect of high flow therapy on hospital length of stay. This finding was in accordance with the latest literature.[13,16,18]

This study found an overall low failure rate for standard and high flow therapy. Even though high flow demonstrated an excellent performance as a rescue therapy for standard care, the failure rate of high flow was elevated. The most likely reason was selection bias because sicker patients were more likely to receive high flow therapy. In addition, the flow rate for high flow therapy was on average lower than the now recommended flow rate of two Litres per minute and kilogram body weight.

This is the first study to describe the relationship between temporal changes of the respiratory severity score and performance of non-invasive ventilation. An ascending trend line of the severity score predicted failure of treatment. It was found to be significantly different between failure and success group. The gradual increase in severity scores might become a useful predictor of impending failure.

In summary, advanced data processing and machine learning provided invaluable insights into the increasing data pool of electronic health records. A data-driven respiratory severity score was developed and partially validated for predicting outcome in infants suffering from acute viral bronchiolitis. In a digitalised healthcare environment, a fully validated respiratory severity score might become a useful tool for effective and efficient management of respiratory support therapy in infants with acute viral bronchiolitis.

References

1.   Meyer AM, Wheeler SB, Weinberger M, Chen RC, Carpenter WR. An overview of methods for comparative effectiveness research. *Semin Radiat Oncol.* 2014;24(1):5-13.
2.   Goss CH, Tefft N. Comparative effectiveness research - what is it and how does one do it? *Paediatric respiratory reviews.* 2013;14(3):152-156.
3.   Florin TA, Plint AC, Zorc JJ. Viral bronchiolitis. *The Lancet.* 2017;389(10065):211-224.
4.   Kyler KE, McCulloh RJ. Current Concepts in the Evaluation and Management of Bronchiolitis. *Infect Dis Clin North Am.* 2018;32(1):35-45.
5.   Drysdale SB, Green CA, Sande CJ. Best practice in the prevention and management of paediatric respiratory syncytial virus infection. *Ther Adv Infect Dis.* 2016;3(2):63-71.
6.   Ralston SL, Lieberthal AS, Meissner HC, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics.* 2014;134(5):e1474-1502.
7.   Hasegawa K, Tsugawa Y, Brown DF, Mansbach JM, Camargo CA, Jr. Trends in bronchiolitis hospitalizations in the United States, 2000-2009. *Pediatrics.* 2013;132(1):28-36.
8.   McNaughten B, Hart C, Shields M. Management of bronchiolitis in infants: key clinical questions. *Paediatrics and Child Health.* 2017;27(7):324-327.
9.   Ducharme-Crevier L, Essouri S, Emeriaud G. Noninvasive ventilation in pediatric intensive care: from a promising to an established therapy, but for whom, when, why, and how? *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies.* 2015;16(5):481-482.
10.   Turnham H, Agbeko RS, Furness J, Pappachan J, Sutcliffe AG, Ramnarayan P. Non-invasive respiratory support for infants with bronchiolitis: a national survey of practice. *BMC pediatrics.* 2017;17(1):20.
11.   Donlan M, Fontela PS, Puligandla PS. Use of continuous positive airway pressure (CPAP) in acute viral bronchiolitis: a systematic review. *Pediatric pulmonology.* 2011;46(8):736-746.
12.   Pedersen MB, Vahlkvist S. Comparison of CPAP and HFNC in Management of Bronchiolitis in Infants and Young Children. *Children (Basel).* 2017;4(4).
13.   Milesi C, Essouri S, Pouyau R, et al. High flow nasal cannula (HFNC) versus nasal continuous positive airway pressure (nCPAP) for the initial respiratory management of acute viral bronchiolitis in young infants: a multicenter randomized controlled trial (TRAMONTANE study). *Intensive care medicine.* 2017;43(2):209-216.
14.   Sinha IP, McBride AK, Smith R, Fernandes RM. CPAP and High-Flow Nasal Cannula Oxygen in Bronchiolitis. *Chest.* 2015;148(3):810-823.
15.   Mace AO, Gibbons J, Schultz A, Knight G, Martin AC. Humidified high-flow nasal cannula oxygen for bronchiolitis: should we go with the flow? *Archives of disease in childhood.* 2018;103(3):303.
16.   Franklin D, Babl FE, Schlapbach LJ, et al. A Randomized Trial of High-Flow Oxygen Therapy in Infants with Bronchiolitis. *The New England journal of medicine.* 2018;378(12):1121-1131.
17.   Riese J, Porter T, Fierce J, Riese A, Richardson T, Alverson BK. Clinical Outcomes of Bronchiolitis After Implementation of a General Ward High Flow Nasal Cannula Guideline. *Hosp Pediatr.* 2017;7(4):197-203.
18.   Kepreotes E, Whitehead B, Attia J, et al. High-flow warm humidified oxygen versus standard low-flow nasal cannula oxygen for moderate bronchiolitis (HFWHO RCT): an open, phase 4, randomised controlled trial. *Lancet.* 2017;389(10072):930-939.
19.   Guillot C, Le Reun C, Behal H, et al. First-line treatment using high-flow nasal cannula for children with severe bronchiolitis: Applicability and risk factors for failure. *Archives de pediatrie : organe officiel de la Societe francaise de pediatrie.* 2018;25(3):213-218.
20.   Korppi M, Tapiainen T. High-flow oxygen therapy is safe and effective in infants with bronchiolitis. *Acta paediatrica (Oslo, Norway : 1992).* 2018.
21.   Ross MK, Wei W, Ohno-Machado L. "Big data" and the electronic health record. *Yearbook of medical informatics.* 2014;9:97-104.
22.   Raman SR, Curtis LH, Temple R, et al. Leveraging electronic health records for clinical research. *American heart journal.* 2018;202:13-19.

23. Sutherland SM, Kaelber DC, Downing NL, Goel VV, Longhurst CA. Electronic Health Record-Enabled Research in Children Using the Electronic Health Record for Clinical Discovery. *Pediatric clinics of North America.* 2016;63(2):251-268.

24. Suresh S. Big Data and Predictive Analytics: Applications in the Care of Children. *Pediatric clinics of North America.* 2016;63(2):357-366.

25. Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. *Clinical research in cardiology : official journal of the German Cardiac Society.* 2017;106(1):1-9.

26. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. *JAMA : the journal of the American Medical Association.* 2018;319(13):1317-1318.

27. Williams JB, Ghosh D, Wetzel RC. Applying Machine Learning to Pediatric Critical Care Data. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies.* 2018.

28. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. *The New England journal of medicine.* 2017;376(26):2507-2509.

29. Lee J, Little TD. A practical guide to propensity score analysis for applied clinical research. *Behav Res Ther.* 2017;98:76-90.

30. Rosenbaum PR. *Design of Observational Studies.* 1st ed. New York: Springer-Verlag; 2010.

31. Burden A, Roche N, Miglio C, et al. An evaluation of exact matching and propensity score methods as applied in a comparative effectiveness study of inhaled corticosteroids in asthma. *Pragmat Obs Res.* 2017;8:15-30.

32. https://clinicaltrials.gov/. Accessed June 2016.

33. http://www.anzctr.org.au/trialSearch.aspx. Accessed June 2016.

34. Bekhof J, Reimink R, Brand PLP. Systematic review: Insufficient validation of clinical scores for the assessment of acute dyspnoea in wheezing children. *Paediatric respiratory reviews.* 2014;15(1):98-112.

35. Lin JA, Madikians A. From bronchiolitis guideline to practice: A critical care perspective. *World J Crit Care Med.* 2015;4(3):152-158.

36. Eggink H, Brand P, Reimink R, Bekhof J. Clinical Scores for Dyspnoea Severity in Children: A Prospective Validation Study. *PloS one.* 2016;11(7):e0157724.

37. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology.* 2007;60(1):34-42.

38. Rodriguez-Martinez CE, Sossa-Briceno MP, Nino G. Systematic review of instruments aimed at evaluating the severity of bronchiolitis. *Paediatric respiratory reviews.* 2017.

39. Marlais M, Evans J, Abrahamson E. Clinical predictors of admission in infants with acute bronchiolitis. *Archives of disease in childhood.* 2011;96(7):648-652.

40. Balaguer M, Alejandre C, Vila D, et al. Bronchiolitis Score of Sant Joan de Deu: BROSJOD Score, validation and usefulness. *Pediatric pulmonology.* 2017;52(4):533-539.

41. van Miert C, McNamara P. Measuring Clinical Severity in Infants with Bronchiolitis. LIBSS-PRO. PhD Thesis. 2015.

42. Riese J, Fierce J, Riese A, Alverson BK. Effect of a Hospital-wide High-Flow Nasal Cannula Protocol on Clinical Outcomes and Resource Utilization of Bronchiolitis Patients Admitted to the PICU. *Hosp Pediatr.* 2015;5(12):613-618.

43. Abboud PA, Roth PJ, Skiles CL, Stolfi A, Rowin ME. Predictors of failure in infants with viral bronchiolitis treated with high-flow, high-humidity nasal cannula therapy*. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies.* 2012;13(6):e343-349.

44. McKiernan C, Chua LC, Visintainer PF, Allen H. High flow nasal cannulae therapy in infants with bronchiolitis. *The Journal of pediatrics.* 2010;156(4):634-638.

45. Evans J, Marlais M, Abrahamson E. Clinical predictors of nasal continuous positive airway pressure requirement in acute bronchiolitis. *Pediatric pulmonology.* 2012;47(4):381-385.

46. Essouri S, Laurent M, Chevret L, et al. Improved clinical and economic outcomes in severe bronchiolitis with pre-emptive nCPAP ventilatory strategy. *Intensive care medicine.* 2014;40(1):84-91.

47.   Ganu SS, Gautam A, Wilkins B, Egan J. Increase in use of non-invasive ventilation for infants with severe bronchiolitis is associated with decline in intubation rates over a decade. *Intensive care medicine.* 2012;38(7):1177-1183.

48.   Mansbach JM, Piedra PA, Stevenson MD, et al. Prospective multicenter study of children with bronchiolitis requiring mechanical ventilation. *Pediatrics.* 2012;130(3):e492-500.

49.   Milési C, Pierre A-F, Deho A, et al. A multicenter randomized controlled trial of a 3-L/kg/min versus 2-L/kg/min high-flow nasal cannula flow rate in young infants with severe viral bronchiolitis (TRAMONTANE 2). *Intensive care medicine.* 2018;44(11):1870-1878.

50.   Chidini G, Piastra M, Marchesi T, et al. Continuous positive airway pressure with helmet versus mask in infants with bronchiolitis: an RCT. *Pediatrics.* 2015;135(4):e868-875.

51.   Bueno Campana M, Olivares Ortiz J, Notario Munoz C, et al. High flow therapy versus hypertonic saline in bronchiolitis: randomised controlled trial. *Archives of disease in childhood.* 2014;99(6):511-515.

52.   Milesi C, Matecki S, Jaber S, et al. 6 cmH2O continuous positive airway pressure versus conventional oxygen therapy in severe viral bronchiolitis: a randomized trial. *Pediatric pulmonology.* 2013;48(1):45-51.

53.   Hilliard TN, Archer N, Laura H, et al. Pilot study of vapotherm oxygen delivery in moderately severe bronchiolitis. *Archives of disease in childhood.* 2012;97(2):182-183.

54.   Kim IK, Phrampus E, Sikes K, et al. Helium-oxygen therapy for infants with bronchiolitis: a randomized controlled trial. *Archives of pediatrics & adolescent medicine.* 2011;165(12):1115-1122.

55.   Essouri S, Durand P, Chevret L, et al. Optimal level of nasal continuous positive airway pressure in severe viral bronchiolitis. *Intensive care medicine.* 2011;37(12):2002-2007.

56.   Mokkink LB, Terwee CB, Patrick DL, et al. COnsensus-based Standards for the selection of health Measurement INstruments. COSMIN checklist manual. *http://wwwcosminnl/indexhtml.* 2012.

57.   Powell CVE. How accurate is the clinical assessment of acute dyspnoea and wheeze in children? *Archives of disease in childhood.* 2015;100(9):813-815.

58.   Wood DW, Downes JJ, Lecks HI. A clinical scoring system for the diagnosis of respiratory failure. Preliminary report on childhood status asthmaticus. *American journal of diseases of children (1960).* 1972;123(3):227-228.

59.   Duarte-Dorado DM, Madero-Orostegui DS, Rodriguez-Martinez CE, Nino G. Validation of a scale to assess the severity of bronchiolitis in a population of hospitalized infants. *The Journal of asthma : official journal of the Association for the Care of Asthma.* 2013;50(10):1056-1061.

60.   Miotto R, Li L, Kidd BA, Dudley JT. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Sci Rep.* 2016;6:26094.

61.   Lasko TA, Denny JC, Levy MA. Computational phenotype discovery using unsupervised feature learning over noisy, sparse, and irregular clinical data. *PloS one.* 2013;8(6):e66341.

62.   Network SIG. Bronchiolitis in children. Quick Reference Guide. Scottish Intercollegiate Guidelines Network. 2006.

63.   Bronchiolitis in children: diagnosis and management. NICE guideline. 2015.

64.   Garcia-Maurino C, Moore-Clingenpeel M, Wallihan R, et al. Discharge Criteria for Bronchiolitis: An Unmet Need. *The Pediatric infectious disease journal.* 2018;37(6):514-519.

65.   Bonafide CP, Brady PW, Keren R, Conway PH, Marsolo K, Daymont C. Development of heart and respiratory rate percentile curves for hospitalized children. *Pediatrics.* 2013;131(4):e1150-1157.

66.   Eytan D, Goodwin AJ, Greer R, Guerguerian AM, Laussen PC. Heart Rate and Blood Pressure Centile Curves and Distributions by Age of Hospitalized Critically Ill Children. *Frontiers in pediatrics.* 2017;5:52.

67.   Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants.pdf>. *BMC pediatrics.* 2013;13.

68.   (CDC). CfDCaP. National Center for Health Statistics. CDC Growth Charts. 2018; https://www.cdc.gov/growthcharts/data_tables.htm. Accessed June 27, 2018.

69.   Golan-Tripto I, Goldbart A, Akel K, Dizitzer Y, Novack V, Tal A. Modified Tal Score: Validated score for prediction of bronchiolitis severity. *Pediatric pulmonology.* 2018;53(6):796-801.

70. McCallum GB, Morris PS, Wilson CC, et al. Severity scoring systems: are they internally valid, reliable and predictive of oxygen use in children with acute bronchiolitis? *Pediatric pulmonology.* 2013;48(8):797-803.

71. Tal A, Bavilski C, Yohai D, Bearman JE, Gorodischer R, Moses SW. Dexamethasone and salbutamol in the treatment of acute wheezing in infants. *Pediatrics.* 1983;71(1):13-18.

72. Caserta MT, Qiu X, Tesini B, et al. Development of a Global Respiratory Severity Score (GRSS) for Respiratory Syncytial Virus Infection in Infants. *The Journal of infectious diseases.* 2016.

73. Brooks CG, Harrison WN, Ralston SL. Association Between Hypertonic Saline and Hospital Length of Stay in Acute Viral Bronchiolitis: A Reanalysis of 2 Meta-analyses. *JAMA pediatrics.* 2016.

74. Sanchez-Pinto LN, Venable LR, Fahrenbach J, Churpek MM. Comparison of variable selection methods for clinical predictive modeling. *International journal of medical informatics.* 2018;116:10-17.

75. Bagherzadeh-Khiabani F, Ramezankhani A, Azizi F, Hadaegh F, Steyerberg EW, Khalili D. A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results. *Journal of clinical epidemiology.* 2016;71:76-85.

76. R-core@R-project.org. Glm - Fitting Generalized Linear Models. R package: stats v3.5.1. 2018; https://www.rdocumentation.org/packages/stats/versions/3.5.1/topics/glm. Accessed Nov 04. 2018.

77. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *2010.* 2010;33(1):22.

78. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. *2010.* 2010;36(11):13.

79. Kapelner A, Bleich J. bartMachine: Machine Learning with Bayesian Additive Regression Trees. *2016.* 2016;70(4):40.

80. Kapelner A, Bleich J. Package 'bartMachine'. An advanced implementation of Bayesian Additive Regression Trees with expanded features for data analysis and visualization. *CRAN Repository.* 2018;Ver. 1.2.4.2.

81. Chen T, He T. Package 'xgboost'. 2018; https://cran.r-project.org/web/packages/xgboost/xgboost.pdf. Accessed Nov 04, 2018.

82. Kapelner A, Bleich J. bartMachine: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software.* 2016;70(4):40.

83. Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. *Ann Appl Stat.* 2010;4(1):266-298.

84. Adam K, Justin B. Prediction with missing data via Bayesian Additive Regression Trees. *Canadian Journal of Statistics.* 2015;43(2):224-239.

85. Bleich J, Kapelner A, George EI, Jensen JT. Variable Selection Inference for Bayesian Additive Regression Trees. *ArXiv e-prints.* 2013.

86. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. *arXiv:160302754.* 2016.

87. Yeo I-K, Johnson RA. A New Family of Power Transformations to Improve Normality or Symmetry. *Biometrika.* 2000;87(4):954-959.

88. Weisberg S. Yeo-johnson power transformations. *Department of Applied Statistics, University of Minnesota.* 2001;2:1-4.

89. Schratz P. Introducing R Package 'oddsratio'. 2017; https://pjs-web.de/post/oddsratio/. Accessed Nov 2018.

90. Mansbach JM. Clinical features can help predict which infants with bronchiolitis will need hospital admission. *The Journal of pediatrics.* 2012;160(1):174-175.

91. Fryzek JP, Martone WJ, Groothuis JR. Trends in chronologic age and infant respiratory syncytial virus hospitalization: An 8-year cohort study. *Advances in Therapy.* 2011;28(3):195-201.

92. Walsh P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R. A validated clinical model to predict the need for admission and length of stay in children with acute bronchiolitis. *European journal of emergency medicine : official journal of the European Society for Emergency Medicine.* 2004;11(5):265-272.

93. Soilly AL, Ferdynus C, Desplanches O, Grimaldi M, Gouyon JB. Paediatric intensive care admissions for respiratory syncytial virus bronchiolitis in France: Results of a retrospective survey and evaluation of the validity of a medical information system programme. *Epidemiology and Infection.* 2012;140(4):608-616.

94.   Rodriguez-Martinez CE, Sossa-Briceno MP, Nino G. Predictors of prolonged length of hospital stay for infants with bronchiolitis. *Journal of investigative medicine : the official publication of the American Federation for Clinical Research.* 2018.

95.   Hervás D, Reina J, Yañez A, del Valle JM, Figuerola J, Hervás JA. Epidemiology of hospitalization for acute bronchiolitis in children: differences between RSV and non-RSV bronchiolitis. *European Journal of Clinical Microbiology & Infectious Diseases.* 2012;31(8):1975-1981.

96.   Jartti T, Aakula M, Mansbach JM, et al. Hospital Length-of-stay Is Associated With Rhinovirus Etiology of Bronchiolitis. *The Pediatric infectious disease journal.* 2014;33(8):829-834.

97.   Alverson B, McCulloh RJ, Dawson-Hahn E, Smitherman SE, Koehn KL. The clinical management of preterm infants with bronchiolitis. *Hosp Pediatr.* 2013;3(3):244-250.

98.   Luo G, Nkoy FL, Gesteland PH, Glasgow TS, Stone BL. A systematic review of predictive modeling for bronchiolitis. *International journal of medical informatics.* 2014;83(10):691-714.

99.   Hasegawa K, Pate BM, Mansbach JM, et al. Risk factors for requiring intensive care among children admitted to ward with bronchiolitis. *Academic pediatrics.* 2015;15(1):77-81.

100.   Parker MJ, Allen U, Stephens D, Lalani A, Schuh S. Predictors of major intervention in infants with bronchiolitis. *Pediatric pulmonology.* 2009;44(4):358-363.

101.   Prodhan P, Sharoor-Karni S, Lin J, Noviski N. Predictors of respiratory failure among previously healthy children with respiratory syncytial virus infection. *The American journal of emergency medicine.* 2011;29(2):168-173.

102.   Wing R, James C, Maranda LS, Armsby CC. Use of high-flow nasal cannula support in the emergency department reduces the need for intubation in pediatric acute respiratory insufficiency. *Pediatric emergency care.* 2012;28(11):1117-1123.

103.   Lazner MR, Basu AP, Klonin H. Non-invasive ventilation for severe bronchiolitis: Analysis and evidence. *Pediatric pulmonology.* 2012;47(9):909-916.

104.   Long E, Babl FE, Duke T. Is there a role for humidified heated high-flow nasal cannula therapy in paediatric emergency departments? *Emergency Medicine Journal.* 2016;33(6):386.

105.   Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate behavioral research.* 2011;46(3):399-424.

106.   Sekhon J. Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software.* 2011;42(7):1-52.

107.   Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software.* 2011;42(8).

108.   Ali MS, Groenwold RH, Klungel OH. Best (but oft-forgotten) practices: propensity score methods in clinical nutrition research. *The American journal of clinical nutrition.* 2016;104(2):247-258.

109.   Randolph JJ, Falbe K, Manueal A, Balloun J. A Step-by-Step Guide to Propensity Score Matching in R. *Practical Assessment, Research & Evaluation.* 2014;19(18):6.

110.   Ergul AB, Calıskan E, Samsa H, et al. Using a high-flow nasal cannula provides superior results to OxyMask delivery in moderate to severe bronchiolitis: a randomized controlled study. *European journal of pediatrics.* 2018;177(8):1299-1307.

111.   Metge P, Grimaldi C, Hassid S, et al. Comparison of a high-flow humidified nasal cannula to nasal continuous positive airway pressure in children with acute bronchiolitis: experience in a pediatric intensive care unit. *European journal of pediatrics.* 2014;173(7):953-958.

112.   Kallappa C, Hufton M, Millen G, Ninan TK. Use of high flow nasal cannula oxygen (HFNCO) in infants with bronchiolitis on a paediatric ward: a 3-year experience. *Archives of disease in childhood.* 2014;99(8):790-791.

113.   Milani GP, Plebani AM, Arturi E, et al. Using a high-flow nasal cannula provided superior results to low-flow oxygen delivery in moderate to severe bronchiolitis. *Acta paediatrica (Oslo, Norway : 1992).* 2016.

114.   Mayfield S, Bogossian F, O'Malley L, Schibler A. High-flow nasal cannula oxygen therapy for infants with bronchiolitis: pilot study. *Journal of paediatrics and child health.* 2014;50(5):373-378.

115.   Kelly GS, Simon HK, Sturm JJ. High-flow nasal cannula use in children with respiratory distress in the emergency department: predicting the need for subsequent intubation. *Pediatric emergency care.* 2013;29(8):888-892.

116.   Betters KA, Gillespie SE, Miller J, Kotzbauer D, Hebbar KB. High flow nasal cannula use outside of the ICU; factors associated with failure. *Pediatric pulmonology.* 2016;52(6):806-812.

117.   Shein SL, Slain KN, Rotta AT. High Flow Nasal Cannula Flow Rates: New Data Worth the Weight. *The Journal of pediatrics.* 2017;189:9-10.

118.   Heikkila P, Sokuri P, Mecklin M, et al. Using high-flow nasal cannulas for infants with bronchiolitis admitted to paediatric wards is safe and feasible. *Acta paediatrica (Oslo, Norway : 1992).* 2018.

119.   Keren R, Shah SS, Srivastava R, et al. Comparative effectiveness of intravenous vs oral antibiotics for postdischarge treatment of acute osteomyelitis in children. *JAMA pediatrics.* 2015;169(2):120-128.

120.   Gupta P, Green JW, Tang X, et al. Comparison of high-frequency oscillatory ventilation and conventional mechanical ventilation in pediatric respiratory failure. *JAMA pediatrics.* 2014;168(3):243-249.

121.   Gupta P, Kneyber M, Essouri S, Rimensberger PC. Letters. Comparison of high-frequency oscillatory ventilation and conventional mechanical ventilation in pediatric respiratory failure. Comment & Response. *JAMA pediatrics.* 2014;168(9):861-863.

122.   Gallego B, Dunn AG, Coiera E. Role of electronic health records in comparative effectiveness research. 2013;2(6).

123.   Ballestero Y, De Pedro J, Portillo N, Martinez-Mugica O, Arana-Arri E, Benito J. Pilot Clinical Trial of High-Flow Oxygen Therapy in Children with Asthma in the Emergency Service. *The Journal of pediatrics.* 2018;194:204-210 e203.

124.   Medina A, del Villar-Guerra P, Modesto i Alapont V. CPAP support should be considered as the first choice in severe bronchiolitis. *European journal of pediatrics.* 2018.

125.   Rubin J, Potes C, Xu-Wilson M, et al. An ensemble boosting model for predicting transfer to the pediatric intensive care unit. *International journal of medical informatics.* 2018;112:15-20.

126.   Ross CE, Harrysson IJ, Goel VV, et al. Integration of Single-Center Data-Driven Vital Sign Parameters into a Modified Pediatric Early Warning System. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies.* 2017;18(5):469-476.

127.   Le Lagadec MD, Dwyer T. Scoping review: The use of early warning systems for the identification of in-hospital patients at risk of deterioration. *Australian critical care : official journal of the Confederation of Australian Critical Care Nurses.* 2017;30(4):211-218.

128.   Lambert V, Matthews A, MacDonell R, Fitzsimons J. Paediatric early warning systems for detecting and responding to clinical deterioration in children: a systematic review. *BMJ Open.* 2017;7(3):e014497.

129.   Birken CS, Parkin PC, Macarthur C. Asthma severity scores for preschoolers displayed weaknesses in reliability, validity, and responsiveness. *Journal of clinical epidemiology.* 2004;57(11):1177-1181.

Index

Apppendix A.   **Items Used for Score**

Numerical Items

Table 29: List of Numerical Items Used for Severity Score

Name of Data Field	Data Type	Description
FiO2	int	Fraction of Inspired Oxygen
Oxygen Administration FIO2	int	FiO$_2$ documented with respiratory support therapy
CBG FIO2	int	FiO$_2$ documented with capillary blood gas
VBG FIO2	int	FiO$_2$ documented with venous blood gas
ABG FIO2	int	FiO$_2$ documented with arterial blood gas
Pulse Oximetry	int	External measurement of oxygen saturation
Heart Rate	int	beats per minute
Heart Rate (bpm) (ICU)	int	beats per minute
Heart Rate, Post Tx	int	beats per minute
Heart Rate, Pre Tx	int	beats per minute
Respiratory Rate	int	breaths per minute
Respiratory Rate (bpm)	int	breaths per minute
Resp. Rate (bpm) (ICU)	int	breaths per minute
Respiratory Rate, Post Tx	int	breaths per minute
Respiratory Rate, Pre Tx	int	breaths per minute
Systolic Blood Pressure	int	mmHg
Arterial Systolic Blood Pressure	int	mmHg
Diastolic Blood Pressure	int	mmHg
Arterial Diastolic Blood Pressure	int	mmHg
Temperature	double	°C
ABG pH	double	pH of arterial blood gas
ABG pO2	int	pO$_2$ of arterial blood gas
CBG pH	double	pH of capillary blood gas
CPG pO2	int	pO$_2$ of capillary blood gas
CBG pCO2	int	pCO$_2$ of capillary blood gas

int = integer, double = floating point value

Text Items with Weight>0

Table 30: List of Text Items Used for Severity Score with Weight Greater Than Zero

ID	Name of Score	Name of Clinical Event	Text items with weight>0	VarDef_ID
21	ssMood	Affect/Mood of Patient	Anxious=2, Fearful=2, Irritable=1, Other: Irritable=1, Other: restless=1	1047
22	ssConscious	Level of Consciousness	Irritable=1, Lethargic=2, Obtunded=1, pt in respiratory distress=2, Restless=1	1524
23	ssSkinColor	Skin Color	Cyanotic=3, desat=3, Dusky=5, Mottled=1, Other: dusky when agitated=3	1944
24	ssRoomAir	Room Air	No=3	1901
37	ssNasalFlaring	Nasal Flaring	Mild=1, Moderate=2, Present=1, Severe=3	1603
38	ssRespType	Respiration Type	Agonal=4, Apnea=3, Bradypnea=2, Controlled with Ventilator=5, Dyspnea=2, Gasping=4, Grunting=3, HFOV or Jet Vent - Unable to Assess=4, increased retractions=2, increased work of breathing=2, Irregular=1, nasal flaring=2, Other: head bobbing=3, Other: WITH NASAL CANNULA=2, Shallow=1, subcostal retraction=1, Tachypnea=2, With Ventilator=5	1876
39	ssRespPEWS	Respiratory PEWS	1. Any assisted ventilation=2, 1. Any supplemental oxygen required to maintain normal =1, 1. Mild retractions=2, 1. O2 > 21% FiO2 required to maintain SpO2=1, 1. Using accessory muscles=3, 2. Moderate retractions=4, 3. Severe retractions and grunting=4	2187
40	ssRespEffort	Respiratory Effort	Increased=1, Mild Distress=2, Moderate Distress=3, Other: Grunting=4, Other: Head bobbing=3, Retractions=1, Severe Distress=4, tachypneic=1	1879
40	ssRespEffort	Respiratory Effort - ED	Grunting=3, Nasal Flaring=2, Other: head bobbing=3, Other: tachypneic=1, Retractions=2	2242
41	ssRespRetraction	Respiratory Retraction	All Muscles Used=4, head bob=3, head bobbing=3, Intercostal=2, Mild=2, Moderate=4, Other: Head bobbing=3, Other: Head bobbing=3, Severe=6, Subcostal=1, Substernal=1, Supraclavicular=3, Suprasternal=3	1887
42	ssUpperAirwaySounds	Upper Airway Sounds	Grunting=2, Other: wheeze=1, Other: wheezing=1, wheezing=1	2095
43	ssBreathSounds	Breath Sounds - Rt Upper	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing=1	1143
43	ssBreathSounds	Breath Sounds - Rt Middle, Post Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1141
43	ssBreathSounds	Breath Sounds - Rt Lower	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing=1	1137

ID	Name of Score	Name of Clinical Event	Text items with weight>0	VarDef_ID
43	ssBreathSounds	Breath Sounds Left - ED	Wheezing=1	2245
43	ssBreathSounds	Breath Sounds - Lt Lower	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing=1	1130
43	ssBreathSounds	Breath Sounds Right - ED	Wheezing=1	2246
43	ssBreathSounds	Breath Sounds - Lt Upper	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing=1	1133
43	ssBreathSounds	Breath Sounds - Rt Middle	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing=1	1140
43	ssBreathSounds	Breath Sounds - RTE	Wheezing and/or rhonchi=1	1146
43	ssBreathSounds	Breath Sounds - Rt Upper, Pre Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1145
43	ssBreathSounds	Breath Sounds - Lt Upper, Pre Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1135
43	ssBreathSounds	Breath Sounds - Lt Lower, Pre Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1132
43	ssBreathSounds	Breath Sounds - Rt Upper, Post Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1144
43	ssBreathSounds	Breath Sounds - Rt Lower, Pre Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1139
43	ssBreathSounds	Breath Sounds Bilateral	Diminished=1, Diminished in Bases=1, Expiratory=1, Inspiratory=1, Wheezing=2	1147
43	ssBreathSounds	Breath Sounds Bilateral - ED	Diminished=1, Wheezing=1	2244
43	ssBreathSounds	Breath Sounds - Bilateral, Post Tx	Diminished=1, Expiratory=1, Expiratory, Wheezing - Diffuse=2, HFOV or Jet Vent - Unable to Assess=4, Inspiratory=1, still wheezing=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1127
43	ssBreathSounds	Breath Sounds - Rt Middle, Pre Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1142
43	ssBreathSounds	Breath Sounds - Bilateral, Pre Tx	Diminished=1, Expiratory=1, HFOV or Jet Vent - Unable to Assess=4, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1128
43	ssBreathSounds	Breath Sounds - Lt Upper, Post Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1134
43	ssBreathSounds	Breath Sounds - Left	Expiratory=1, Wheezing=1	1129
43	ssBreathSounds	Breath Sounds - Lt Lower, Post Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1131
43	ssBreathSounds	Breath Sounds - Right	Expiratory=1, Wheezing=1	1136
43	ssBreathSounds	Breath Sounds - Rt Lower, Post Tx	Diminished=1, Expiratory=1, Inspiratory=1, Wheezing - Diffuse=2, Wheezing - Focal=1	1138

Text Items with Weight=0

Table 31: Text Items Used for Severity Score with Weight Equal Zero

ID	Name of Score	Name of Clinical Event	Text items with weight=0	VarDef_ID
21	ssMood	Affect/Mood of Patient	Acting Out, Awake, Crying, Interactive, Other: kicking, Other: moving., Playful, Quiet, Sleeping, Uncooperative, Upbeat	1047
22	ssConscious	Level of Consciousness	Agitated, Alert, Arousable, Asleep, Awake, Combative, Crying, Drowsy, Other: consolable, Other: coughing, Other: patient being held by mother, Playful, Quiet, Responsive	1524
23	ssSkinColor	Skin Color	Flushed, Jaundiced, Normal for Ethnicity, Pale, Pink	1944
24	ssRoomAir	Room Air	Room air, Yes	1901
37	ssNasalFlaring	Nasal Flaring	Absent, None Noted	1603
38	ssRespType	Respiration Type	Regular	1876
39	ssRespPEWS	Respiratory PEWS	0. No retractions, 0. Within normal parameters for age	2187
40	ssRespEffort	Respiratory Effort	No Distress, None, Normal	1879
40	ssRespEffort	Respiratory Effort - ED	Comfortable	2242
41	ssRespRetraction	Respiratory Retraction	None, Other: none noted	1887
42	ssUpperAirwaySounds	Upper Airway Sounds	Clear	2095
43	ssBreathSounds	Breath Sounds - Rt Lower, Post Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered	1138
43	ssBreathSounds	Breath Sounds - Bilateral, Pre Tx	Anterior, Bowel Sounds, bs fairly clear, Clear, Coarse, Coarse Crackles, Diffuse, Diminished in Bases, Fine Crackles, Moist Crackles, Pleural Rub, Posterior, Rhonchi, Scattered, Squeaks	1128
43	ssBreathSounds	Breath Sounds - Rt Middle, Post Tx	Anterior, Clear, Coarse, Coarse Crackles, Fine Crackles, Moist Crackles, Rhonchi, Scattered	1141
43	ssBreathSounds	Breath Sounds - Left	Clear	1129
43	ssBreathSounds	Breath Sounds Left - ED	Clear	2245
43	ssBreathSounds	Breath Sounds - Right	Clear	1136

ID	Name of Score	Name of Clinical Event	Text items with weight=0	VarDef_ID
43	ssBreathSounds	Breath Sounds Right - ED	Clear	2246
43	ssBreathSounds	Breath Sounds - Rt Upper	Anterior, Clear, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Focal, Moist Crackles, Posterior, Rhonchi, Scattered, Squeaks	1143
43	ssBreathSounds	Breath Sounds - Lt Upper	Anterior, Clear, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Focal, Moist Crackles, Pleural Rub, Posterior, Rhonchi, Scattered, Squeaks	1133
43	ssBreathSounds	Breath Sounds - Rt Lower	Anterior, Clear, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Focal, Moist Crackles, Pleural Rub, Posterior, Rhonchi, Scattered, Squeaks	1137
43	ssBreathSounds	Breath Sounds - Rt Middle	Absent, Anterior, Clear, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Focal, Moist Crackles, Pleural Rub, Posterior, Rhonchi, Scattered, Squeaks	1140
43	ssBreathSounds	Breath Sounds - Lt Lower	Anterior, Clear, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Focal, Moist Crackles, Posterior, Rhonchi, Scattered, Squeaks	1130
43	ssBreathSounds	Breath Sounds - RTE	Clear to auscultation, Crackles in bases, Decrease Bilaterally, Decrease Unilaterally	1146
43	ssBreathSounds	Breath Sounds - Rt Upper, Pre Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered, Squeaks	1145
43	ssBreathSounds	Breath Sounds - Lt Upper, Pre Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Posterior, Rhonchi, Scattered, Squeaks	1135
43	ssBreathSounds	Breath Sounds - Lt Lower, Pre Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered	1132
43	ssBreathSounds	Breath Sounds - Rt Upper, Post Tx	Anterior, Bowel Sounds, Clear, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Moist Crackles, Posterior, Rhonchi, Scattered	1144
43	ssBreathSounds	Breath Sounds - Rt Lower, Pre Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered, Squeaks	1139
43	ssBreathSounds	Breath Sounds Bilateral - ED	Clear	2244
43	ssBreathSounds	Breath Sounds - Rt Middle, Pre Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered, Squeaks	1142
43	ssBreathSounds	Breath Sounds - Lt Upper, Post Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered	1134
43	ssBreathSounds	Breath Sounds Bilateral	Before Treatment, Clear & Equal, Coarse, Coarse Crackles, Diffuse, Fine Crackles, Focal, Moist Crackles, Other: stider, Rhonchi, Scattered	1147
43	ssBreathSounds	Breath Sounds - Lt Lower, Post Tx	Anterior, Clear, Coarse, Coarse Crackles, Diminished in Bases, Fine Crackles, Moist Crackles, Rhonchi, Scattered, Squeaks	1131
43	ssBreathSounds	Breath Sounds - Bilateral, Post Tx	94% sleeping, Anterior, Clear, Coarse, Coarse Crackles, Diffuse, Diminished in Bases, even after suctioning, Fine Crackles, Moist Crackles, patient crying., Pleural Rub, post neb tx, post Neb Tx., Posterior, remains diminished., Rhonchi, Scattered, Squeaks	1127

Apppendix B.   **Types of Clinical Events**

Numerical

See Apppendix A: Items Used for Score, Table 29, page 118

Categorical

The original raw dataset contained 9385 different types of clinical events.

Final table with clinical events contains 2,687,505 records containing 563 different types of events. The table below lists the most common (occurrence>10,000) event types.

Table 32: List of the most common event types (occurence >10,000)

Cnt	EVENT_CD	C_EVENT_DISP	Result Example
105656	688374	Pulse Oximetry	97
98703	688367	Heart Rate	187
98055	671599	Respiratory Rate	47
94123	1004652	Administration Information	0
56812	708296	Level of Consciousness	Alert, Awake
53215	671596	Systolic Blood Pressure	123
53211	671597	Diastolic Blood Pressure	57
52672	736249	Type of Oxygen Administration	With Ventilator
51517	708918	Skin Color	Normal for Ethnicity, Pink
51217	702384	Respiration Type	Tachypnea
51197	702392	Respiratory Effort	Increased
47840	702422	Nasal Flaring	Absent
44995	702466	Chest Expansion	Equal
41820	788168	Liter Flow Per Minute	0.5
41127	77301884	Breath Sounds Bilateral	Clear, Equal
40760	688867	Temperature	37
40109	687918	Temperature Route	Axillary
37818	702578	Upper Airway Sounds	Clear
36931	702641	Cough	Absent
35163	752905	FiO2	21
33979	820857	Affect/Mood of Patient	Irritable
33028	732198	Extremity Temperature	Warm
31953	702401	Respiratory Retraction	Mild
31790	731085	FLACC Primary Pain Cry	1. Moans or Whimpers; Occasional Complaint
31789	731079	FLACC Primary Pain Legs	0. Normal Position or Relaxed
31779	731076	FLACC Primary Pain Face	1. Occasional Grimace or Frown, Withdrawn, Disinterested
31776	731082	FLACC Primary Pain Activity	0. Lying Quietly, Normal Position, Moves Easily
31760	731088	FLACC Primary Pain Consolability	0. Content, Relaxed
31673	731045	FLACC Pain Intensity	2
30083	93961964	Behavioral Assessment	Agitation
28138	88180382	Cuff MAP	67
26463	106351222	24hr Urine Output (Weight-based)	2.9
24314	688860	Mean Arterial Pressure	79
24205	819569	Oral Formula Intake 1	120
24028	20375774	Cardio Respiratory Monitor	Off
22723	55026317	Room Air	Room air
21819	702660	Suction Site	Oropharynx
21748	702673	Secretion Amount	Small
20171	102968568	Secretion Characteristics	White, Thick
19329	774096	Trunk Temperature	Warm
19147	93966654	Cardiac Rhythm	Sinus Tachycardia
18951	106351077	1hr Urine Output (Weight-based)	0
18936	95587268	Pulse Pressure	58
18207	133000393	Behavior PEWS	0. Playing, appropriate
18204	133009314	Respiratory PEWS	0. Within normal parameters. No retractions
18200	133008082	Cardiovascular PEWS	0. Pink color. Capillary refill <2 seconds
18186	133009818	Medical History PEWS	0. N/A
18132	133010567	PEWS Score	0
17814	903441	fentanyl	5
17471	93967600	Peripheral Capillary Refill	2

Cnt	EVENT_CD	C_EVENT_DISP	Result Example
17420	93967857	Central Capillary Refill	2
16877	719010	Radial Pulse, Left	2+ Normal
16857	719007	Radial Pulse, Right	2+ Normal
16240	702654	Cough Upon Request	Strong
15973	708930	Mucous Membranes	Moist
15830	773938	Brachial Pulse, Left	2+ Normal
15820	773935	Brachial Pulse, Right	2+ Normal
14247	719019	Capillary Filling	Normal (<2 Seconds)
14247	98983714	Eye Opening Response	4. Spontaneous
14219	98983720	Best Motor Response	6. Moves spontaneously/purposely (infant)
14207	98983734	Best Verbal Response	4. Irritable Cry (infant)
14140	34222778	FLACC Pain Tool	Yes
14135	98983700	Total GCS	14
14079	103060996	Respiratory Procedures	Suction, Pulse Ox Probe Site Change
13430	731846	Dextrose 5% in Water	15
10979	115033859	Pediatric Care Team	Team 4
10440	735410	Respiratory Treatment Response	Partially Effective
10373	736230	Respiratory Treatment Given Via	Face mask

The final set of clinical events used for data analysis contains 62 clinical event types.

Table 33: List of the final set of clinical events used for data analysis (n=62)

ID	EVENT_CD	VarName	VarDim	VarType	
1010	743723	ABG FIO2	%	NUM	integer
2273	743735	CBG FIO2	%	NUM	integer
3240	743746	VBG FIO2	%	NUM	integer
1358	752905	FiO2	%	NUM	integer
1654	783255	Oxygen Administration FIO2	%	NUM	integer
1836	688374	Pulse Oximetry	%	NUM	integer
104	688367	Heart Rate	bpm	NUM	integer
5975	818065	Heart Rate (bpm) (ICU)	bpm	NUM	integer
1414	102969341	Heart Rate, Post Tx	bpm	NUM	integer
1415	102969398	Heart Rate, Pre Tx	bpm	NUM	integer
8666	818066	Resp. Rate (bpm) (ICU)	bpm	NUM	integer
110	671599	Respiratory Rate	bpm	NUM	integer
1885	102969337	Respiratory Rate, Post Tx	bpm	NUM	integer
1886	102969405	Respiratory Rate, Pre Tx	bpm	NUM	integer
3095	343374007	Respiratory Rate (bpm)	bpm	NUM	integer
1081	824485	Arterial Diastolic Blood Pressure	mm HG	NUM	integer
1083	824481	Arterial Systolic Blood Pressure	mm HG	NUM	integer
1277	671597	Diastolic Blood Pressure	mm HG	NUM	integer
2008	671596	Systolic Blood Pressure	mm HG	NUM	integer
1018	743718	ABG pO2	mm HG	NUM	integer
1178	743729	CBG pCO2	mm HG	NUM	integer
1180	743730	CBG pO2	mm HG	NUM	integer
1016	743716	ABG pH	Unit	NUM	float
1179	743728	CBG pH	Unit	NUM	float
2017	688867	Temperature	deg C	NUM	float
1047	820857	Affect/Mood of Patient	(null)	TXT	
1095	93961964	Behavioral Assessment	(null)	TXT	
1127	102969320	Breath Sounds - Bilateral, Post Tx	(null)	TXT	
1128	102969381	Breath Sounds - Bilateral, Pre Tx	(null)	TXT	
1129	103060525	Breath Sounds - Left	(null)	TXT	
1130	55030866	Breath Sounds - Lt Lower	(null)	TXT	
1131	102969267	Breath Sounds - Lt Lower, Post Tx	(null)	TXT	
1132	102969348	Breath Sounds - Lt Lower, Pre Tx	(null)	TXT	
1133	55030499	Breath Sounds - Lt Upper	(null)	TXT	
1134	102969281	Breath Sounds - Lt Upper, Post Tx	(null)	TXT	
1135	102969352	Breath Sounds - Lt Upper, Pre Tx	(null)	TXT	
1136	103060537	Breath Sounds - Right	(null)	TXT	
1137	55030494	Breath Sounds - Rt Lower	(null)	TXT	
1138	102969295	Breath Sounds - Rt Lower, Post Tx	(null)	TXT	
1139	102969358	Breath Sounds - Rt Lower, Pre Tx	(null)	TXT	
1140	55030551	Breath Sounds - Rt Middle	(null)	TXT	
1141	102969308	Breath Sounds - Rt Middle, Post Tx	(null)	TXT	
1142	102969369	Breath Sounds - Rt Middle, Pre Tx	(null)	TXT	
1143	55030481	Breath Sounds - Rt Upper	(null)	TXT	
1144	102969313	Breath Sounds - Rt Upper, Post Tx	(null)	TXT	
1145	102969375	Breath Sounds - Rt Upper, Pre Tx	(null)	TXT	
1146	34064043	Breath Sounds - RTE	(null)	TXT	
1147	77301884	Breath Sounds Bilateral	(null)	TXT	

ID	EVENT_CD	VarName	VarDim	VarType
2244	383062763	Breath Sounds Bilateral - ED	(null)	TXT
2245	383062812	Breath Sounds Left - ED	(null)	TXT
2246	383062856	Breath Sounds Right - ED	(null)	TXT
1524	708296	Level of Consciousness	(null)	TXT
1579	708930	Mucous Membranes	(null)	TXT
1603	702422	Nasal Flaring	(null)	TXT
1876	702384	Respiration Type	(null)	TXT
1879	702392	Respiratory Effort	(null)	TXT
2242	383062676	Respiratory Effort - ED	(null)	TXT
2187	133009314	Respiratory PEWS	(null)	TXT
1887	702401	Respiratory Retraction	(null)	TXT
1901	55026317	Room Air	(null)	TXT
1944	708918	Skin Color	(null)	TXT
2095	702578	Upper Airway Sounds	(null)	TXT

Distribution of Clinical Events

Table 34: List of used clinical events orderder by the usage in percent

Name of Data Field	Units	Data Type	Count	Percentage
Pulse Oximetry	%	NUM	42258	9.28%
Heart Rate	bpm	NUM	38252	8.40%
Respiratory Rate	bpm	NUM	37999	8.34%
Level of Consciousness		TXT	24304	5.34%
Respiration Type		TXT	22411	4.92%
Diastolic Blood Pressure	mm HG	NUM	22032	4.84%
Systolic Blood Pressure	mm HG	NUM	22029	4.84%
Respiratory Effort		TXT	21888	4.81%
Skin Color		TXT	21620	4.75%
Affect/Mood of Patient		TXT	21509	4.72%
Nasal Flaring		TXT	20937	4.60%
Upper Airway Sounds		TXT	18977	4.17%
Temperature	deg C	NUM	18066	3.97%
Breath Sounds Bilateral		TXT	17992	3.95%
Room Air		TXT	15656	3.44%
Respiratory Retraction		TXT	14380	3.16%
Respiratory PEWS		TXT	11990	2.63%
FiO2	%	NUM	7973	1.75%
Respiratory Rate, Pre Tx	bpm	NUM	5508	1.21%
Heart Rate, Pre Tx	bpm	NUM	5504	1.21%
Respiratory Rate, Post Tx	bpm	NUM	5422	1.19%
Heart Rate, Post Tx	bpm	NUM	5407	1.19%
Breath Sounds - Bilateral, Pre Tx		TXT	5340	1.17%
Breath Sounds - Bilateral, Post Tx		TXT	5276	1.16%
Breath Sounds - Right		TXT	4112	0.90%
Breath Sounds - Left		TXT	4094	0.90%
Breath Sounds - Rt Upper		TXT	1764	0.39%
Breath Sounds - Rt Lower		TXT	1735	0.38%
Breath Sounds - Lt Lower		TXT	1725	0.38%
Breath Sounds - Lt Upper		TXT	1723	0.38%
Breath Sounds - Rt Middle		TXT	1684	0.37%
CBG pH	Unit	NUM	1043	0.23%
CBG pCO2	mm HG	NUM	1041	0.23%
CBG pO2	mm HG	NUM	1040	0.23%
CBG FIO2	%	TXT	344	0.08%
Breath Sounds - RTE		TXT	202	0.04%
Breath Sounds - Rt Upper, Pre Tx		TXT	191	0.04%
Breath Sounds - Lt Upper, Pre Tx		TXT	166	0.04%
Respiratory Effort - ED		TXT	165	0.04%
Breath Sounds - Lt Lower, Pre Tx		TXT	159	0.03%
Breath Sounds - Rt Lower, Pre Tx		TXT	159	0.03%
Breath Sounds - Rt Middle, Pre Tx		TXT	152	0.03%

Name of Data Field	Units	Data Type	Count	Percentage
Breath Sounds Bilateral - ED		TXT	151	0.03%
Breath Sounds - Rt Upper, Post Tx		TXT	149	0.03%
Breath Sounds - Lt Upper, Post Tx		TXT	136	0.03%
Breath Sounds - Lt Lower, Post Tx		TXT	133	0.03%
Breath Sounds - Rt Lower, Post Tx		TXT	121	0.03%
Breath Sounds - Rt Middle, Post Tx		TXT	119	0.03%
Breath Sounds Left - ED		TXT	117	0.03%
Breath Sounds Right - ED		TXT	115	0.03%
Oxygen Administration FIO2	%	NUM	32	0.01%
VBG FIO2	%	NUM	29	0.01%
ABG pH	Unit	NUM	18	0.00%
ABG pO2	mm HG	NUM	18	0.00%
ABG FIO2	%	NUM	14	0.00%

Name of Data Field	Units	Data Type	Count	Percentage
Breath Sounds Bilateral - ED		TXT	151	0.03%
Breath Sounds - Rt Upper, Post Tx		TXT	149	0.03%
Breath Sounds - Lt Upper, Post Tx		TXT	136	0.03%

Apppendix C.   **Entity Relationship Diagram (ERD)**

pers

«column»
*PK P_ID: INT
* DOB: DATE
DMin: BIGINT
Sex: CHAR(1)
* SN: INT
GA_Days: INT
Apgar1: INT
Apgar5: INT
Apgar10: INT
BirthWeight: DOUBLE
BirthLength: DOUBLE
BirthWeight_z: DOUBLE
BirthLength_z: DOUBLE
CreatedOn: DATETIME
ModifiedOn: DATETIME

«PK»
+ PK_pers(INT)

«index»
+ idx_pers_DMin(BIGINT)
+ idx_pers_SN(INT)

+PK_pers   +fk_diag_P_ID

+PK_pers

+PK_pers

+fk_cevnt_P_ID

diag

«column»
*PK D_ID: INT
* N_ID: INT
*FK P_ID: INT
*FK E_ID: INT
SOURCE_IDENTIFIER: VARCHAR(50)
D_DIAG_TYPE_DISP: VARCHAR(50)
DIAG_DT_TM: DATETIME
UPDT_DT_TM: DATETIME
BEG_EFFECTIVE_DT_TM: DATETIME
END_EFFECTIVE_DT_TM: DATETIME
DIAG_Min: INT
UPDT_Min: INT
BEG_EFFECTIVE_Min: INT
END_EFFECTIVE_Min: INT
D_CLASSIFICATION_DISP: VARCHAR(50)
D_CLINICAL_SERVICE_DISP: VARCHAR(50)
DIAGNOSIS_DISPLAY: VARCHAR(100)
DIAGNOSIS_GROUP: INT
CLINICAL_DIAG_PRIORITY: INT
DIAG_TYPE_CD: INT
ICD_CM: TINYINT

«PK»
+ PK_diag(INT)

«index»
+ idx_diag_D_DIAG_TYPE_DISP(VARCHAR)
+ idx_diag_E_ID(INT)
+ idx_diag_N_ID(INT)
+ idx_diag_P_ID(INT)
+ idx_diag_SOURCE_IDENTIFIER(VARCHAR)

«FK»
+ fk_diag_E_ID(INT)
+ fk_diag_P_ID(INT)

+fk_diag_E_ID

+PK_enctr

cevnt

«column»
*PK C_ID: BIGINT
*FK P_ID: INT
*FK E_ID: INT
* EVENT_CD: INT
C_EVENT_DISP: VARCHAR(100)
RESULT_VAL: TEXT
C_RESULT_UNITS_DISP: VARCHAR(50)
C_EVENT_CLASS_DISP: VARCHAR(50)
* EventMon: INT
EventDay: VARCHAR(12)
* EventDayMin: INT
* eventAgeDays: INT
eventResultInt: INT
eventResultReal: DOUBLE
EVENT_START_Min: INT
EVENT_END_Min: INT
PERFORMED_Min: INT
VERIFIED_Min: INT
C_ENTRY_MODE_DISP: VARCHAR(50)
PARENT_EVENT_ID: BIGINT
C_SOURCE_DISP: VARCHAR(50)

«PK»
+ PK_cevnt(BIGINT)

«index»
+ idx_cevnt_eventAgeDays(INT)
+ idx_cevnt_EventDay(VARCHAR)
+ idx_cevnt_EventMon(INT)
+ idx_cevnt_Event_CD(INT)
+ idx_cevnt_EVENT_END_Min(INT)
+ idx_cevnt_EVENT_START_Min(INT)
+ idx_cevnt_E_ID(INT)
+ idx_cevnt_PARENT_EVENT_ID(BIGINT)
+ idx_cevnt_PERFORMED_Min(INT)
+ idx_cevnt_P_ID(INT)
+ idx_cevnt_RESULT_VAL(TEXT)
+ idx_cevnt_VERIFIED_Min(INT)

«FK»
+ fk_cevnt_E_ID(INT)
+ fk_cevnt_P_ID(INT)

+fk_enctr_P_ID

+fk_cevnt_E_ID

+PK_enctr

enctr

«column»
*PK E_ID: INT
*FK P_ID: INT
REG_DT_TM: DATETIME
DISCH_DT_TM: DATETIME
* REG_Min: INT
* DISCH_Min: INT
* AGEREG_DAYS: DOUBLE
* AGEREG_MIN: INT
* LOS_DAYS: DOUBLE
* LOS_MIN: INT
E_ENCNTR_CLASS_DISP: VARCHAR(50)
E_ENCNTR_STATUS_DISP: VARCHAR(50)
E_ENCNTR_TYPE_DISP: VARCHAR(50)
E_DISCH_DISPOSITION_DISP: VARCHAR(50)
REASON_FOR_VISIT: VARCHAR(255)
E_MED_SERVICE_DISP: VARCHAR(155)
E_LOCATION_DISP: VARCHAR(155)
ENCNTR_CLASS_CD: INT
ENCNTR_STATUS_CD: INT
ENCNTR_TYPE_CD: INT
DISCH_DISPOSITION_CD: INT
LOCATION_CD: INT

«PK»
+ PK_enctr(INT)

«index»
+ idx_enctr_AGEREG_DAYS(DOUBLE)
+ idx_enctr_AGEREG_MIN(INT)
+ idx_enctr_DISCH_Min(INT)
+ idx_enctr_E_DISCH_DISPOSITION_DISP(VARCHAR)
+ idx_enctr_LOS_DAYS(DOUBLE)
+ idx_enctr_LOS_MIN(INT)
+ idx_enctr_REG_Min(INT)

«FK»
+ fk_enctr_P_ID(INT)

Apppendix D.   **Cerner Command Language (CCL)**

Examples of CCL-code to demonstrate the approach of step-wise data extraction.

Find Diagnoses and Patients

```
SELECT
      D.PERSON_ID
      , P.BIRTH_DT_TM "MM/DD/YYYY HH:MM;;D"
      , P_SEX_DISP = UAR_GET_CODE_DISPLAY(P.SEX_CD)
      , N.SOURCE_IDENTIFIER
      , N.SOURCE_STRING
      , D.NOMENCLATURE_ID
      , D.DIAG_TYPE_CD
      , D_DIAG_TYPE_DISP = UAR_GET_CODE_DISPLAY(D.DIAG_TYPE_CD)
      , D.ENCNTR_ID
      , D.DIAG_DT_TM "MM/DD/YYYY HH:MM;;D"
      , D.UPDT_DT_TM "MM/DD/YYYY HH:MM;;D"
      , D_CLASSIFICATION_DISP = UAR_GET_CODE_DISPLAY(D.CLASSIFICATION_CD)
      , D_CLINICAL_SERVICE_DISP = UAR_GET_CODE_DISPLAY(D.CLINICAL_SERVICE_CD)
      , D.DIAGNOSIS_DISPLAY
      , D.DIAGNOSIS_GROUP
      , D.CLINICAL_DIAG_PRIORITY
      , D.DIAGNOSIS_ID
FROM
      NOMENCLATURE N
      , DIAGNOSIS D
      , PERSON P
PLAN N
      where (n.source_identifier = "466.1*" or n.source_identifier = "J21.*")
      and n.active_ind = 1

      join d
      where d.nomenclature_id = n.nomenclature_id
      and D.ACTIVE_IND = 1

      join p
      where p.PERSON_ID = d.PERSON_ID
      and P.ACTIVE_IND = 1

WITH MAXREC = 15000, NOCOUNTER, TIME = 300
```

Find Encounters

```
SELECT
      P.PERSON_ID
      , mrn = pa.alias
      , P.BIRTH_DT_TM "MM/DD/YYYY HH:MM;;D"
      , P_SEX_DISP = UAR_GET_CODE_DISPLAY(P.SEX_CD)
      , E.ENCNTR_ID
      , AgeREG_days = DATETIMEDIFF(E.REG_DT_TM,P.BIRTH_DT_TM,1) ; days
      , AgeREG_min = DATETIMEDIFF(E.REG_DT_TM,P.BIRTH_DT_TM,4) ; min
      , E.REG_DT_TM "MM/DD/YYYY HH:MM;;D"
      , E.DISCH_DT_TM "MM/DD/YYYY HH:MM;;D"
      , LOS_days = DATETIMEDIFF(E.DISCH_DT_TM,E.REG_DT_TM,1) ; days
      , LOS_min = DATETIMEDIFF(E.DISCH_DT_TM,E.REG_DT_TM,4) ; minutes
      , E_ENCNTR_CLASS_DISP = UAR_GET_CODE_DISPLAY(E.ENCNTR_CLASS_CD)
      , E_ENCNTR_STATUS_DISP = UAR_GET_CODE_DISPLAY(E.ENCNTR_STATUS_CD)
      , E_ENCNTR_TYPE_DISP = UAR_GET_CODE_DISPLAY(E.ENCNTR_TYPE_CD)
      , E_DISCH_DISPOSITION_DISP = UAR_GET_CODE_DISPLAY(E.DISCH_DISPOSITION_CD)
      , E.ENCNTR_CLASS_CD
      , E.ENCNTR_STATUS_CD
      , E.ENCNTR_TYPE_CD
      , E.DISCH_DISPOSITION_CD
FROM
      NOMENCLATURE   N
```

```
      , DIAGNOSIS   D
      , PERSON     P
      , ENCOUNTER    E
      , person_alias   pa
PLAN N
      where (n.source_identifier = "466.1*" or n.source_identifier = "J21.*")
      and n.active_ind = 1

      join d
      where d.nomenclature_id = n.nomenclature_id
      and D.ACTIVE_IND = 1

      join p
      where p.PERSON_ID = d.PERSON_ID
      and P.ACTIVE_IND = 1

      join pa
      where pa.person_id = p.person_id
      AND pa.PERSON_ALIAS_TYPE_CD = 10

      join e
      where E.PERSON_ID = D.PERSON_ID
      and e.ENCNTR_CLASS_CD NOT IN (319457, 319458) ; NOT Outpatient, Preadmit
      and DATETIMEDIFF(E.DISCH_DT_TM,E.REG_DT_TM)>=2 ;LOS in days >2
      and DATETIMEDIFF(E.REG_DT_TM,P.BIRTH_DT_TM)<366

WITH MAXREC = 99000, NOCOUNTER, TIME = 300
```

Find clinical events for Each Patient And Encounter

As per worksheet "DATA" 14,438,356 events extracted,
Save in subfolder "Clinical Events; 466.1, J21.; AgeReg 1-365 days, LOS 2-x days"
under 01.csv - 218.csv (3.47 GB)

```
SELECT
      C.CLINICAL_EVENT_ID
      , C.EVENT_CD
      , C.PERSON_ID
      , P.BIRTH_DT_TM "MM/DD/YYYY HH:MM;;D"
      , P_SEX_DISP = UAR_GET_CODE_DISPLAY(P.SEX_CD)
      , E.ENCNTR_ID
      , E.REG_DT_TM "MM/DD/YYYY HH:MM;;D"
      , E.DISCH_DT_TM "MM/DD/YYYY HH:MM;;D"
      , C_EVENT_DISP = UAR_GET_CODE_DISPLAY(C.EVENT_CD)
      , C_EVENT_CLASS_DISP = UAR_GET_CODE_DISPLAY(C.EVENT_CLASS_CD)
      , C.EVENT_START_DT_TM "MM/DD/YYYY HH:MM;;D"
      , C.EVENT_END_DT_TM "MM/DD/YYYY HH:MM;;D"
      , C.PERFORMED_DT_TM "MM/DD/YYYY HH:MM;;D"
      , C.VERIFIED_DT_TM "MM/DD/YYYY HH:MM;;D"
      , C.RESULT_VAL
      , C_RESULT_UNITS_DISP = UAR_GET_CODE_DISPLAY(C.RESULT_UNITS_CD)
      , C_ENTRY_MODE_DISP = UAR_GET_CODE_DISPLAY(C.ENTRY_MODE_CD)
      , C.PARENT_EVENT_ID
      , C_SOURCE_DISP = UAR_GET_CODE_DISPLAY(C.SOURCE_CD)

FROM
      PERSON    P
      , ENCOUNTER    E
      , CLINICAL_EVENT    C

PLAN P
      where p.ACTIVE_IND = 1

      join e
      where e.PERSON_ID = p.PERSON_ID
      and e.ACTIVE_IND = 1
```

```
        join c
        where c.ENCNTR_ID = e.ENCNTR_ID
        and c.RESULT_STATUS_CD IN (25) ; include only "Auth (Verified)"


    and (

(p.PERSON_ID = 11684765 and e.ENCNTR_ID = 31601027)
or (p.PERSON_ID = 11682761 and e.ENCNTR_ID = 31593519)
; etc.

WITH MAXREC = 145000, NOCOUNTER, TIME = 300
```

Apppendix E.   **Final Dataset**

Description of Final Dataset

URL: http://ckcdata.com/R-code/

```
url_all <- "http://ckcdata.com/R-code/pivot_ss_all2.csv"
url_sumSS <- "http://ckcdata.com/R-code/avb_sumSS2.csv"
url_sumSS180 <- "http://ckcdata.com/R-code/avb_sumSS180.csv"

url_ssNIV <- "http://ckcdata.com/R-code/avb_ssNIV7.csv"
url_ssNIV_all <- "http://ckcdata.com/R-code/avb_ssNIV_all7.csv"

url_ssFirst12hrs <- "http://ckcdata.com/R-code/sumSS_First_12hrs.csv"

md_all <- read.csv(url_all, header = TRUE)
md_summSS <- read.csv(url_sumSS, header = TRUE)
md_sumSS180 <- read.csv(url_sumSS180, header = TRUE)
md_ssNIV <- read.csv(url_ssNIV, header = TRUE)
md_ssNIV_all <- read.csv(url_ssNIV_all, header = TRUE)
md_ssFirst12hrs <- read.csv(url_ssFirst12hrs, header = TRUE)

gNC <- 0
gHFNC <- 1
gNCesc <- 2
gHFNCesc <- 3
```

All Cases - md_all

md	Obs.	Var.	NC	HFNC	CMV	NIV_selected	Successful NC NC=1, HFNC=0	Escalated NC NC=1, HFNC=1	Successful HF HFNC=1, CMV=0	Escalated HF NC=0, HFNC=1, CMV=1
md_all	627	40	540	157	20	NC=505, HFNC=122	470	35 NIV_selected=NC	103 NIV_selected=HFNC	18
md_sumSS	592	16	505	122	19	NA	470	35	103	18

pivot_ss_all

avb_sumSS180

avb_ssNIV

avb_ssNIV_all

sumSS_First_12hrs

Apppendix F.   **Example of R Code**

Correlation between Severity Score and Length of Stay

```
---
title: "5.4-Correlation-severity-score-Length-of-Stay-v1.1"
author: "Dr Christoph Camphausen"
date: "Nov 18, 2018"
output: html_document
---

## Check Model Performance to predict tLOS
tLOS~sumSS+zWt+AGEREG_DAYS+RSV+virusMetapneumo+Male
Severity score measured at the time of the treatment decision
to apply either standard nasal oxygen or high flow nasal oxygen

md <- md_ssNIV_all

Transorm <- 1

if (Transorm==1) {
  coVarYJ <- c("zWt", "AGEREG_DAYS","tLOS")
  md_  <- md[,coVarYJ]

  # 6. Yeo-Johnson Transform
  # Another power-transform like the Box-Cox transform,
  # but it supports raw values that are equal to zero and negative.
  lambdaEstimates <- preProcess(md_, method=c("YeoJohnson"))
  mdYJ <- predict(lambdaEstimates, md_)
  md$zWt <- mdYJ$zWt
  md$AGEREG_DAYS <- mdYJ$AGEREG_DAYS
  md$tLOS <- mdYJ$tLOS

  md$sumSS <- log(md$sumSS+1)
}



md1 <- md
covar1 <- c("sumSS","zWt","AGEREG_DAYS","RSV", "virusMetapneumo","Male")

mFormula <- getFF("tLOS",covar1)
(sFormula <- getFF_str("tLOS",covar1))

(nmd <- nrow(md1))
(nCoVar <- length(covar1))

maintitle <- paste0("Model Performance of ",nCoVar," covariates to predict tLOS \n",
                    sFormula," | n=",nmd)
maintitle

# 10-fold cross validation, repeated 10 times
trainControl <- trainControl(method="repeatedcv", number=10, repeats=3)
metric <- "RMSE"

# LM
set.seed(7)
fit.lm <- train(mFormula, data=md1, method="lm", metric=metric,
                trControl=trainControl)
# GLM
set.seed(7)
fit.glm <- train(mFormula, data=md1, method="glm",
                 metric=metric, trControl=trainControl)

# Generalized Linear Model with Stepwise Feature Selection
set.seed(7)
fit.glmStep <- train(mFormula, data=md1, method="glmStepAIC", metric=metric,
                     trControl=trainControl)

# GLMNET
set.seed(7)
fit.glmnet <- train(mFormula, data=md1, method="glmnet", metric=metric,
                    trControl=trainControl)
# SVM
set.seed(7)
fit.svm <- train(mFormula, data=md1, method="svmRadial",
                 metric=metric,
                 trControl=trainControl)

# KNN
```

```
set.seed(7)
fit.knn <- train(mFormula, data=md1, method="knn",
                 metric=metric,
                 trControl=trainControl)

# Compare algorithms
resultsRegression <- resamples(list(LM=fit.lm,GLM=fit.glm, GLMNET=fit.glmnet,StepAIC=fit.glmStep,
                                    SVM=fit.svm, KNN=fit.knn ))

summary(resultsRegression)
dotplot(resultsRegression, main = maintitle)
```

Apppendix G.   **Bekhof Quality Criteria for Dyspnoea Scores**

Table 35: Measurement Instrument: Validity

Validity			
Face Validity	Qualitative judgement if the score is a good measurement of dyspnoea.[129]	+	At least 3 of the following items were part of the score: 1. respiratory and/or heart rate, 2. oxygen saturation or cyanosis, 3. work of breathing, retractions or use of muscles or dyspnoea, 4. wheezing or auscultatory findings, 5. mental status.
		±	2 of the above mentioned items
		-	1 item
Content validity*	Appropriate representation of the concept dyspnoea by the items in the score (i.e., clear description of development process of the score). [11]	+	Clear description is provided for measurement aim, target population and item selection and item reduction
		?	Potential methodological shortcomings
		-	No clear description
Construct validity*	Extent to which the score relates to other measures, consistent with theoretically derived pre-specified hypotheses concerning dyspnoea [Kirschner 1985, Terwee 2007]	+	Specific hypotheses were formulated and at least 75% of the results are in correspondence with these hypotheses in subgroups of at least 50 patients.
		?	Less than 50 patients OR potential methodological shortcomings or no MIC
		-	Less than 75% of the hypotheses are confirmed
		0	No information
Criterion-concurrent validity*	Criterion validity refers to the extent to which a score relates to the gold standard of the phenomenon. Because a gold standard of dyspnoea is unavailable, this is replaced by concurrent validity, the degree of agreement with other measurements of dyspnoea. [11]	+	Valid comparison (oxygen saturation, laboratory findings or pulmonary function tests) and correlation >0.70
		?	Doubts about gold standard
		-	Correlation < 0.70
		0	No information

Table 36: Measurement Instrument: Reliability

Reliability			
Measure-ment Error*	Absolute measurement error, usually expressed as smallest detectable change (SDC), i.e. the smallest within-person change in score which can be interpreted as real change above measurement error [11]	+	MIC>SDC or LOA<MIC [10]
		?	Potential methodological shortcomings or no MIC
		-	SDC or LOA ≥ MIC
		0	No information
Inter-Observer Reliability*	Degree to which different users obtain the same result when using the score on the same patients at the same time	+	ICC or weighted kappa >0.70 in at least 50 patients
		?	Pearson correlation >0.70, or < 50 patients
		-	ICC or kappa <0,70
		0	No information
Intra-Observer Reliability*	Similarity of results when the score is repeated by the same user on the same patient under similar conditions	+	ICC or weighted kappa >0.70 in at least 50 patients [10]
		?	Pearson correlation >0.70, or < 50 patients
		-	ICC or kappa <0,70
		0	No information
Internal Consistency*	Correlation between items of the score [10,11]	+	Factor analysis performed and Cronbach's alfa 0.70-0.95
		?	No factor analysis OR potential methodological shortcomings
		-	Cronbach's alfa < 0.70 or > 0.95
		0	No information
Responsive-ness*	Ability of the score to detect change in time [9]	+	Guyatts's RR >1.96 or AUC ≥ 0.70
		?	Potential methodological shortcomings
		-	RR ≤ 1.96 of AUC < 0.70
		0	No information

Table 37: Measurement Instrument: Utility

Utility			
Suitability	Suitability for use in children	+	No invasive techniques or items which may be difficult to obtain in young children (e.g. pulsus paradoxus, information on speech not specified for infants).
		±	As in + with information on speech specified for infants
		-	Use of invasive techniques or items which may be difficult to obtain in young children
Age Span	Coverage of the entire paediatric age span	+	Evaluated from infancy (<2 years)
		-	Evaluated in a smaller age span
Ease of Scoring	Complexity of scoring system	+	<4 categories per item [47]
		±	4 categories per item
		-	> 4 categories per item or complex calculations needed
Auscultation Skills	Feasibility in clinical practice by different health care providers	+	No auscultation skills required
		±	no complex auscultation skills required (no inspiratory:expiratory ratio).
		-	complex auscultation skills required
Floor or Ceiling Effect*	Unequal distribution of score results [11]	+	< 15% of patients with lowest or highest possible score in at least 50 patients
		?	Potential methodological shortcomings or < 50 patients
		-	< 15% of patients with lowest or highest possible score
		0	No information
Interpretability*	Clinical meaningfulness	+	Mean scores and SD given in at least 4 relevant subgroups and MIC determined
		?	Potential methodological shortcomings or < 4 subgroups or no MIC determined
		0	No information
			Sum

"+ positive rating; ± indeterminate; - negative rating; ? unclear or potential methodological shortcomings; 0 no information available (potential) methodological shortcomings = description of design or methods of the study not clear, or study group < 50 persons.
* Items together form the Terwee checklist [11].
SDC smallest detectable change; MIC minimal important change; LOA limits of agreement; ICC intraclass correlation coefficient; SD standard deviation; RR Guyatt's responsiveness ratio; AUC area under curve of the receiver operating curve."

MACQUARIE University
SYDNEY·AUSTRALIA

2 May 2017

Dear Dr Gallego Luxan

Reference No: 5201700411

Title: *Use of Prediction Models to Optimize Respiratory Support Therapy of Infants with Acute Viral Bronchiolitis: Retrospective Observational Study using Machine-Learning Techniques at a large Tertiary Center*

Thank you for submitting the above application for ethical and scientific review. Your application was considered by the Macquarie University Human Research Ethics Committee (HREC (Medical Sciences)).

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted at:

- Macquarie University

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007 – Updated May 2015) (the *National Statement*).

Standard Conditions of Approval:

1. Continuing compliance with the requirements of the *National Statement,* which is available at the following website:

http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research

2. This approval is valid for five (5) years, subject to the submission of annual reports. Please submit your reports on the anniversary of the approval for this protocol.

3. All adverse events, including events which might affect the continued ethical and scientific acceptability of the project, must be reported to the HREC within 72 hours.

4. Proposed changes to the protocol and associated documents must be submitted to the Committee for approval before implementation.

It is the responsibility of the Chief investigator to retain a copy of all documentation related to this project and to forward a copy of this approval letter to all personnel listed on the project.

Should you have any queries regarding your project, please contact the Ethics Secretariat on 9850 4194 or by email ethics.secretariat@mq.edu.au

The HREC (Medical Sciences) Terms of Reference and Standard Operating Procedures are available from the Research Office website at:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics

The HREC (Medical Sciences) wishes you every success in your research.

Yours sincerely

Professor Tony Eyers
Chair, Macquarie University Human Research Ethics Committee (Medical Sciences)

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice*.

Details of this approval are as follows:

Approval Date: 27 April 2017

The following documentation has been reviewed and approved by the HREC (Medical Sciences):

Documents reviewed	Version no.	Date
Macquarie University Human Research Ethics Committees Prior Review Form (PREF)	N/A	Received 10 Apr 2017

Documents Noted	Version no.	Date
Notice of IRB Exemption – with stipulations (Children's Hospital Los Angeles Institutional Review Board)	N/A	13 Apr 2017
IStar Application (Children's Hospital Los Angeles)	1.3	12 Apr 2017
Research Protocol	2.3	11 Apr 2017
Data Collection Sheet	2.3	11 Apr 2017
Data Use Agreement- Protected Health Information	N/A	Signed 24 Feb 2017

*If the document has no version date listed one will be created for you. Please ensure the footer of these documents are updated to include this version date to ensure ongoing version control.

Email on this page removed from Open Access version as it may contain sensitive/confidential content.