Can RSV-Associated Hospitalization in the First Year of Life be Predicted at Birth Among Infants Born at 32-35 Weeks Gestation?

by

Venessa M J Ryan

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
November 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF COMMUNITY HEALTH AND EPIDEMIOLOGY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Can RSV-Associated Hospitalization in the First Year of Life be Predicted at Birth Among Infants Born at 32-35 Weeks Gestation?" by Venessa MJ Ryan in partial fulfilment of the requirements for the degree of Master of Science.

Dated: November 21, 2012

Supervisor: _____

Readers: _____

_____

DALHOUSIE UNIVERSITY

DATE: November 21, 2012

AUTHOR: Venessa Ryan

TITLE: Can RSV-Associated Hospitalization in the First Year of Life be Predicted at Birth Among Infants Born at 32-35 Weeks Gestation?

DEPARTMENT OR SCHOOL: Department of Community Health and Epidemiology

DEGREE: MSc CONVOCATION: May YEAR: 2013

_____

Signature of Author

DEDICATION PAGE

This thesis is dedicated to my parents, Gary and Monique, who have been my constant source of encouragement and have always supported my endeavours over the years. They have taught me that with determination, discipline and hard work, anything is possible.

This work is dedicated to my sisters, Natalie and Christine, and my grandparents, Ola and Estelle, who inspire me with their enthusiasm and never ending “joie de vivre”. Also, to the memory of my grandparents, Hazel and John, who both passed on a love of their family, and are still constant reminders of the simple pleasures in life.

I am particularly grateful for the immeasurable support and patient encouragement from Fabian, without his love and guidance this work would not have been made possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This retrospective cohort study examined risk factors associated with RSV-associated hospitalization (RSV-H) among infants born 32 to 35 weeks gestational age in Nova Scotia. Results were used to develop a clinical instrument (RSV-H scoring tool) that would discriminate between infants at high, moderate and low risk for RSV-H. Identifying the highest-risk infants, (using baseline information to predict RSV-H in the first year of life), would help target cost effective prophylaxis by Palivizumab (Pz), an expensive RSV-specific monoclonal antibody. Significant risk factors, determined by multivariate logistics, included infants born in December, January or February, passive household smoke exposure and household crowding. The scoring tool produced similar RSV-H post-test probabilities (3.1% pre-test probability) between risk groups (5.5% vs. 5.8%) and was unable to target highest risk infants. The tool could be used as an educational guideline for health professionals, outlining the importance of significant risk factors for RSV-H to parents and caregivers.

LIST OF ABBREVIATIONS AND SYMBOLS USED

$\chi 2$	Chi-squared
<	Less than
≤	Less than or equal to
>	Greater than
≥	Greater than or equal to
95% CI	95% Confidence Interval
AAP	American Academy of Pediatrics
AUC	Area under the curve
BPD	Bronchopulmonary Dysplasia
CA	Chronological age
CDC	Centers for Disease Control and Prevention
CHD	Congenital Heart Disease
CIHI	Canadian Institute for Health Information
CLD	Chronic Lung Disease
CPS	Canadian Paediatric Society
df	Degree of freedom
ETS	Environmental tobacco smoke
GA	Gestational age
HCN	Health card number
ICD-CM	International Classification of Disease Clinical Modification code
LMP	Last menstrual period
LR	Likelihood ratio
LR+	Positive likelihood ratio
LR-	Negative likelihood ratio
LRTI	Lower respiratory tract infection
N	Quantity
PPRoM	Premature preterm rupture of membranes
PPV	Positive predictive value
NPV	Negative predictive value

NS	Nova Scotia
NSAPD	Nova Scotia Atlee Perinatal Database
NSPWG	Nova Scotia Palivizumab Working Group
OR	Odds Ratio
p	P-value
Pz	Palivizumab
RCP	The Reproductive Care Program of Nova Scotia
ROC	Receiver Operator Characteristic
RR	Relative Risk
RSV	Respiratory Syncytial Virus
RSV-H	Respiratory Syncytial Virus-associated Hospitalization
SES	Socioeconomic status
wGA	Weeks gestational age

ACKNOWLEDGEMENTS

I would like to express my very great appreciation to Dr. Joanne Langley, my research supervisor, for her valuable and constructive suggestions during the planning and development of this research work. I would like to express my deep gratitude to my advisory committee Drs. Pantelis Andreou and Linda Dodds for their patient guidance, enthusiastic encouragement and useful critiques of this research work.

Assistance provided by the members of the Joint Data Access Committee through the Reproductive Care Program was greatly appreciated, in particular Dr. Alexander Allen, whose knowledge and advice throughout the ethics submission and approval process was appreciated. I also wish to acknowledge the help provided by staff at the IWK Health Centre research services.

I wish to thank various people for their contribution to this project; research analyst John Fahey of the Reproductive Care Program, for his help on linking the dataset, and his shared knowledge about the dataset variables; Brian Hoyt, for ensuring the secure transfer and extraction of the data for my analysis; database analyst Yan Wang, and staff at the Population Health Research Unit for providing me with access to the dataset, and sharing their knowledge on statistics.

Dr. Victoria Allen has been the ideal mentor, providing me with very valuable research opportunities and has further inspired my intellectual curiosity in epidemiology, and for that I am forever indebted.

My special thanks are extended to staff of the Community Health and Epidemiology Department who ensured a seamless experience, and to the faculty, whose passion for health education and knowledge was so evidently displayed throughout my degree. I thank the students in CH&E for making my Master's experience memorable.

Lastly, I would like to thank my family and friends for providing endless inspiration and support, both for the duration of my master's degree and beyond.

CHAPTER 1: INTRODUCTION

Human Respiratory Syncytial Virus (RSV) is a negative stranded nonsegmented RNA virus classified in the Pneumovirinae subfamily of the Paramyxoviridae family [1]. The virus often causes upper respiratory infections (such as "colds"), and it is a major cause of lower respiratory tract infections (LRTI) (such as bronchiolitis and pneumonia) leading to hospital admissions during infancy and childhood. RSV is highly contagious and by their second year, most infants will have been infected at least once [2]. Although the majority of RSV infections resolve uneventfully in otherwise healthy children, certain high-risk populations may develop severe and sometimes fatal LRTI.

Children who are at increased risk for RSV-associated hospitalization (RSV-H) include those with premature birth or underlying cardiac or lung disease [3, 4]. RSV related illnesses account for approximately 75,000-125,000 hospitalizations per year in the United States [5], while in Canada, this accounts for an estimated 5,800-12,000 hospitalizations annually [6]. In the USA alone, RSV-H costs exceed $400 million (US) annually [5, 7, 8], while a prospective cohort study conducted estimated the annual cost of RSV-associated illness in Canada at nearly $18 million (1993 US dollars) [6].

Due to the contagious nature of RSV, common hygienic practices such as cleaning contaminated surfaces, washing hands frequently and avoiding direct contact or contaminated surfaces infected with the virus can effectively prevent infection and spread of the virus. The only proven medical prophylaxis is by Palivizumab SYNAGIS® (Pz), a monoclonal antibody preparation, given through intramuscular injection in pediatric patients at high risk of RSV-associated complications. It is given in monthly intramuscular injections (up to a maximum of 5) based on the eligibility criteria of the infant. Safety and efficacy were established in infants with chronic lung disease (CLD)/bronchopulmonary dysplasia (BPD), infants with a history of prematurity (≤ 35 wGA), and children with hemodynamically significant congenital heart disease (CHD) [9, 10]

Because of the high cost of palivizumab (~$CDN 5000/child/year) [11], it is prudent to use this medication in the population in which it will be most effective. Systematic literature reviews have been conducted in order to identify studies examining the cost effectiveness of Pz versus no prophylaxis for RSV in children. While deemed cost ineffective for the general population, the systematic reviews demonstrated cost effectiveness in high-risk subpopulations. The reviews concluded that prophylaxis tended to be most cost effective in populations with specific risk factors, including premature infants < or = 32 weeks' gestational age, and infants or children aged < 2 years with CLD or CHD [12, 13]. Similar results were found in a study for the 32-35 wGA subgroup conducted in Spain [14].

Use of palivizumab in Canada is usually directed by a territorial or provincial program, which often only permits the use of Pz in selected infants at highest risk. In Nova Scotia (NS), the Provincial Blood Coordinating Program (NSPBCP) convened a provincial working group of medical specialists representing stakeholders from institutions. The working group reviews RSV prophylaxis cost and utilization data for NS, as well assist in the designing of a utilization management program for RSV immunoprophylaxis products. The committee meets to review data from the previous RSV season and to revise considerations for the NS Provincial RSV Prophylaxis Guidelines[15].

Current NS guidelines for indications of Pz use include children <24 months of age with BPD/CLD, premature infants (defined as born at 32 weeks (+ 0 days) or earlier gestation) with or without BPD/CLD, and young children with congenital heart disease (CHD). The guidelines allow consideration for moderately premature infants (32 1/7 weeks to 35 6/7 weeks gestation) based on collected risk factors, and “other” children considered to be at high risk for contracting RSV (e.g. immunodeficiency or those with severe chronic lung disease not due to prematurity)[15].

There is evidence that moderately premature infants are at increased risk for RSV-H [16-19]. Although Pz prophylaxis has been approved by the Food and Drug

Administration (FDA) for use in premature infants > 32 weeks gestation, most jurisdictions only fund programs that deliver Pz to infants ≤ 32 weeks gestation. There is a high cost associated with providing prophylaxis in moderately premature infants who comprise up to 5% of the birth cohort. Efforts have been made to identify a sub-population of infants ≥ 32 – 35 or 36 wGA who are at highest risk.

Evidence demonstrates an association between RSV-associated severe LRTI and the following risk factors: month of birth (during RSV season), young chronological age (CA) at the beginning of RSV season, infants living with school-aged siblings, child care attendance, absence or short duration of breastfeeding, exposure to tobacco smoke during pregnancy or after birth, household crowding, male gender, low birth weight, ethnicity (Aboriginal), family socioeconomic status (SES), low level of maternal RSV antibody levels in the infant, maternal age, multiple births and family history of allergies/respiratory disease [11, 11, 13, 17, 20-34]. Many provinces have subsequently revised their eligibility criteria to allow moderately premature infants to receive Pz based on the presence of these identified risk factors.

The aims of the current study are to quantify the following:

1. To assess and identify risk factors present at birth that are associated with a higher likelihood of RSV-H among Nova Scotian infants 32 to 35 completed wGA.

2. To develop a model based on significant risk factors and to create a tool that will assist health care professionals in assessing which infants are at highest risk of RSV-H.

Identifying risk factors that are significantly associated with RSV-H would provide health professionals with readily available information on baseline variables that best identify high-risk infants. By developing a model based on these significant risk factors, the specificity and sensitivity of such a tool could support the evidence base recommendations for Pz prophylaxis. Additionally, it would add value to preventive interventions at the IWK Health Centre and across NS, targeting infants 32 to 35 completed wGA.

CHAPTER 2: BACKGROUND

2.1 EPIDEMIOLOGY OF RESPIRATORY SYNCYTIAL VIRUS

Respiratory Syncytial Virus is usually spread by droplets, direct contact or fomites. The incubation period ranges from 2 to 8 days; 4 to 6 days is most common [35]. RSV is responsible for annual epidemics throughout the world. Globally, it is estimated that RSV causes about 34 million episodes of acute lower respiratory infections in children younger than five years, resulting in approximately 3.4 million hospitalizations per year and it is responsible for an estimated 160,000 deaths annually worldwide; however, incidence and mortality can vary substantially from year to year in any one location [36]. In the northern hemisphere, these usually occur between November and April, with a peak in January or February. In the southern hemisphere, wintertime epidemics occur between May and September, with a peak in May, June, or July. In tropical and semitropical climates, the seasonal outbreaks usually are associated with the rainy season [7, 35]. More specifically, in Canada, RSV season usually begins between November and January and persists for four to five months [6, 37, 38].

The Centers for Disease Control and Prevention (CDC) consider RSV to be the most common cause of bronchiolitis and pneumonia in children under 1 year of age [39], with 20% to 30% developing lower respiratory tract disease with their first infection [40, 41]. Overall, out of all children infected, 1-2% will be hospitalized. In children who are hospitalized with RSV LRTI, the mortality rate is less than 1%; however, in select groups of high-risk patients (infants with CLD related to prematurity, CHD, or prematurity), this disease may have a mortality rate as high as 3-5% [7, 42].

2.2 POPULATIONS AT HIGHEST RISK FOR RSV-ASSOCIATED SEVERE LOWER RESPIRATORY TRACT INFECTION

Consistent evidence shows that children with certain medical conditions, specifically CLD, CHD, and prematurity, are at an increased risk for RSV LRTI compared with the general infant population. Unsurprisingly, these higher-risk infants have been consistently identified as the primary risk groups considered for prophylaxis[3,4].

Chronic lung disease is a general term for long-term respiratory problems in new born babies. It is also known as bronchopulmonary dysplasia (BPD). CLD results from lung injury to newborns who must use a mechanical ventilator and extra oxygen for breathing, or were born premature (whose lungs are not fully developed at birth). The lungs of premature babies are fragile and are easily damaged. With injury, the tissues inside the lungs become damaged and cause difficulty breathing and increases oxygen needs. Infants who have had CLD are at greater risk for repeated respiratory infections, such as RSV that require a hospital stay [43]. Studies have also shown an increased need for mechanical ventilation, increased mortality rates (3.4%), and prolonged lengths of stay in both general and intensive care unit admission days in children with CLD compared with otherwise healthy children [44, 45].

Congenital heart disease describes a number of different problems affecting the heart and is the most common type of birth defect. The disease is often divided into two types: cyanotic (blue discoloration caused by a relative lack of oxygen (e.g. Tetralogy of Fallot) and non-cyanotic (e.g.Ventricular Septal Defect, Atrial Septal defect, Patent Ductus Arteriosus). These problems may occur alone or together. They can also be a part of various genetic and chromosomal syndromes such as Down syndrome, trisomy 13 and Turner syndrome [46]. Patients with CHD with RSV experience higher rates of hospital admissions, are more likely to receive mechanical ventilation and have more lengthy hospital stays compared to infants without CHD [19, 47, 48].

Gestational age (GA) is the time measured from the first day of the woman's last menstrual cycle to the current date, and it is measured in weeks. An infant is considered premature if born before 37 weeks gestation. These infants are vulnerable to a variety of illnesses, have smaller, immature airways because they are deprived of the last few weeks of development in utero. Of importance with respect to RSV, the rate of lung maturation is greatest during the third trimester of pregnancy. Interrupted pulmonary development prior to 36 weeks reveals smaller lung volumes, smaller lung surface area, whereas air space walls have an increased wall thickness, in comparison to term infants (those born after 37 weeks gestation). Furthermore, the immature immune system results in incomplete transfer of maternal antibodies [49]. As a result, postnatal exposure to viruses results in increased susceptibility to LRTI [50-52].

Studies indicate that prematurity is a consistently proven risk marker for RSV LRTI [11, 16, 17, 20, 21, 53-57]. While RSV LRTI is generally recognized as posing a serious risk to early premature infants (≤ 32 weeks GA), all premature infants are in some form at higher risk for RSV infection. In a cumulative analysis of 24 studies, levels of morbidity were, at times, comparable between infants born at 32 to 36 wGA and in infants born at earlier gestation [18]. Studies comparing moderate preterm infants with no underlying medical conditions consistently revealed an RSV-H rate similar to that observed in infants born 29 to 33 weeks [16-19]. Mounting evidence suggests that outcomes in these moderate preterm infants are less than optimal, and in some instances, these infants may be equally at risk for RSV-associated severe LRTI as infants born at ≤32 weeks GA [11, 21, 53, 54].

2.3 PREVENTION OF RESPIRATORY SYNCYTIAL VIRUS

The safety and efficacy of Pz were assessed in a randomized, double-blind, placebo controlled multicenter study (IMpact-RSV Trial) of RSV disease prophylaxis among high-risk pediatric patients. Throughout the study, RSV hospitalizations occurred among 53 of 500 (10.6%) patients in the placebo group and 48 of 1,002 (4.8%) patients in the Pz group; a 55% reduction ($p<0.001$). Specifically, the reduction of RSV-H was observed in premature infants born 32 to 35 wGA throughout the course of the study, (9.8% RSV-H among the placebo group vs. 2.0% RSV-H among the Pz recipients; 80% reduction, $p=0.002$)[9]. These findings led to palivizumab's approval by the FDA in June 1998 [58, 59].

In North America, both the American Academy of Paediatrics (AAP) and the Canadian Paediatric Society (CPS) Infectious Diseases and Immunization Committee provide recommendations for administering Pz for RSV prophylaxis based on their evaluation of evidence. The AAP's most recent recommendations for use of Pz was issued in 2009 [60]. In October 2011, the CPS released updated recommendations for administering Pz for RSV prophylaxis (Table 2-1)[61].

2.4 RISK FACTORS ASSOCIATED WITH RSV-ASSOCIATED HOSPITALIZATION IN PREMATURE INFANTS

In recent years, attempts have been made to identify independent risk factors associated with RSV-H in moderately premature infants. RSV-H has been generally defined in the literature as one or combining the following definitions:

1) The hospital admission for a respiratory tract infection episode proved to be RSV by a confirmed laboratory test;

2) The hospital admission for a respiratory tract infection by International Classification of Disease Clinical Modification code (ICD-CM codes) assigned by health record coders based on review of the patient's health record;

3) The hospital admission for a respiratory tract infection by ICD-CM code using hospital discharge information [11, 20-22, 48, 52, 53, 62-64].

The following host and environmental related risk factors have been demonstrated in the literature as being associated with RSV-H; young gestational age (GA < 37 weeks), birth shortly before or early after onset of RSV season (birth month), CA at the beginning of RSV season, multiple births (twins/triplets), male sex, low family SES and parental education, crowded living conditions, exposure to maternal smoking and indoor smoke pollution, small for GA (percentile), low birth weight (grams), family history of asthma or allergies, day care attendance, older siblings in school or day-care, and absence or short duration of breastfeeding [11, 19-21, 32, 49, 55, 56, 65-68]. Other studies have demonstrated an association between the mode of delivery, maternal age and low maternal RSV antibody levels[13, 17, 23, 24, 32-34].

Given the high cost of Pz prophylaxis, particular attention has been given in identifying risk factors that could predict RSV-H and be used to develop and validate risk assessment within moderately premature infants (discussed in section 2.5). Two recent studies of infants born 33-35 wGA identified additional risk factors associated with RSV-H. The Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) study conducted in 16 regions across Canada during 2 successive RSV seasons in which 1860 infants, including those with underlying congestive and/or lung disease, were followed for a mean of 104 days. The Canadian study found that day care attendance was the most significant independent risk factor. Factors that most Likely may lead to development of RSV-related respiratory Infection and subsequent hospital administration among Premature infants born 33 to 35 weeks gestational age study (FLIP) study, compared premature infants hospitalized for RSV with matched controls (n=186 cases and 371 controls). The study concluded that birth < 10 weeks before the start of RSV

season (July-December) was the most significant independent risk factor. In comparison to the PICNIC study, the FLIP study showed no statistically significant increase in risk of hospitalization with birth weight, male gender, child-care attendance, or exposure to tobacco smoke at home. The same researchers published a follow-up to the FLIP trial that followed 5,441 children over 2 RSV seasons among premature infants 32-35 wGA. The aim of the FLIP-2 study was to validate the results of the previous epidemiologic FLIP study during two consecutive RSV seasons. The FLIP-2 study showed that only three independent risk factors–absolute CA ≤10 weeks at the start of the RSV season, presence of school-age siblings or day care attendance and smoking during pregnancy–were found to significantly increase the risk of RSV-H. Tables 2-2 and 2-3 provide summaries of four studies that have demonstrated independent risk factors related to RSV LRTI hospitalization for moderately premature infants [11, 20-22]. A summary of significant risk factors for each study can be seen in Table 2-4.

The differences in results between the Canadian and Spanish studies can be attributed to the difference in study designs. The prospective PICNIC and FLIP-2 studies included the enrolment of all premature infants who fulfilled the inclusion criteria. Whereas the FLIP study had a case-control design, which matched cases and controls. By matching variables, there is always the possibility that matching for known factors also indirectly matches for unknown related factors, and may lead to spurious results. For example, the lack of statistical significance for smokers in the household with the FLIP dataset is likely due to the smaller sample size and the potential confounding by matching variables in the case-control design [11, 69]. The differences in study design also indicate variability in defining the RSV-H group of infants (cases), and the non-hospitalized group (controls) between the studies. By selecting cases, a number of RSV-H may have been negelected in the case-control design. Additionally, the authors of the FLIP-2 study acknowledge a relatively high loss to follow-up of 12% for infants who met the inclusion criteria. The results of the study therefore may have been limited by a selection bias introduced in the methods attributable to difficulties contacting families (changes in follow up, language barrier).

The impact of each risk factor on an individual may vary between settings. This may include geographical location, severity of RSV epidemics, environmental characteristics, diagnostic testing, disease management, access to care and differences in health care practices such as diagnostic testing and disease management. This may account for differences in quantifying the size of the overall effect of each risk factor in relation to RSV severe infection.

2.5 DEVELOPMENT AND VALIDATION OF RISK SCORING TOOLS TO PREDICT RSV-ASSOCIATED HOSPITALIZATION

Presented below are a brief explanation of the European model and a further description of the Canadian risk scoring tool.

The European RSV risk study group developed a predictive model to aid the targeting of prophylaxis throughout Europe for those infants born 33-35 wGA at risk of RSV-H [64], based on risk factors from the Spanish FLIP study. The final model consisted of seven risk factors: birth ±10 weeks of start of season, birth weight, breast fed ≤2 months, number of siblings ≥2 years, number of family members with atopy, number of family members with wheeze, and gender[63].

The Canadian RSV risk scoring tool was developed using factors associated with RSV hospitalization in 1,758 infants born at 33-35 wGA at 16 Canadian centres followed prospectively from birth through the first RSV season [54]. The final model is based on seven risk factors, each with an assigned score. Scores are categorized as low (0-48 points), moderate (49 -64 points) and high (> 65 points). Each risk factor is identified as being either present or absent, with a positive answer being assigned a pre-determined score between 10 and 25. Table 2-5 provides a summary of each risk factor and its assigned score. The Canadian risk scoring tool was retrospectively validated using a study among Spanish infants at the same gestational age [20, 54]. Provinces have started to use this tool or other risk-based strategies to identify eligibility for Pz among premature infants. The PICNIC tool's ability to predict hospitalization was recently assessed and

found to be insufficiently sensitive and specific, in a cohort of Nova Scotian infants (personal communication J. Langley).

Stakeholders are continuously seeking new research-based evidence in order to assist decision making for judicious monoclonal antibody prophylaxis. We hypothesize that by collecting local data on baseline variables in NS infants, RSV-H before 1 year of age can be predicted from risk factors present at birth in infants 32 to 35 completed wGA.

Table 2-1: Current Recommendations for RSV Prophylaxis in North America

Risk Group	AAP recommendation[70]	CPS recommendation[61]
Hemodynamically Significant Chronic lung disease (CLD) of infancy	Consider infants ≤24 months of age who receive medical therapy (supplemental oxygen, bronchodilator, diuretic or chronic corticosteroid therapy) within 6 months before the start of the RSV season; five doses.	Consider infants ≤24 months of age at the start of RSV season; receive up to five doses.
Hemodynamically Significant Congenital heart disease (CHD)	Consider infants ≤24 months of age; five doses. Decisions in children with CHD should be made on the basis of the degree of physiologic cardiovascular compromise.	Consider infants ≤24 months of age at the start of RSV season; receive up to five doses.
≤28 weeks GA	During first RSV season (≤12 months of age) Receive up to five doses.	Receive up to five doses.
29–32 weeks GA	29 weeks, 0 days through to 31 weeks, 6 days. Major **risk factors to consider include gestational age and chronologic age at the start of the RSV season**. May benefit most from prophylaxis if younger than 6 months of age at the start of the RSV season**.**	29 weeks, 0 days through to 31 weeks, 6 days who are younger than six months of age at the start of the RSV season; receive up to five doses.
32–35 weeks GA	32 weeks, 0 days through to 34 weeks 6 days; <3 months of age at start of RSV season; three doses maximum (stop dosing at ≥90 days of age**); one of two risk factors is present (childcare attendance and sibling(s) <5 years of age).**	32 weeks 0 days to 35 weeks 6 days' GA. A panel of experts should be convened in each province or territory to establish a policy for these infants. American Academy of Pediatrics (AAP) criteria or the Canadian risk-scoring tool may be used to select eligible infants.

Table 2-2: Comparison of Four Studies of Risk Factors for RSV-Associated Hospitalization in Moderately Premature Infants Without Other Risk Factors

Study Details	PICNIC[11]	FLIP[20]	FLIP-2[21]	Osservatorio VRS[22]
Study design	Multicenter, nationwide, prospective cohort study	Multicenter, nationwide, case-control study	Multicenter, nationwide, prospective cohort study	Multicenter, nationwide, case-control study
Country and study population	Canada 33 (+ 0 days) to 35 (+ 6 days) wGA	Spain 33 (+ 0 days) to 35 (+ 6 days) wGA	Spain 32 (+ 1 days) to 35 (+ 0 days) wGA	Italy Preterm and full-term children categorized as < 33 wGA, **33-35 wGA** and ≥ 36 wGA
Aim of the study	To identify risk factors associated with hospitalization for RSV infection	To identify risk factors that most likely may lead to development of RSV-related respiratory infection and subsequent hospital admission	To validate the risk factors found in the FLIP case-control study regarding hospitalization as a result of RSV infection	To identify risk factors associated with higher likelihood to acquire RSV-induced LRTI in children with symptoms severe enough to lead to hospital admission
Definition of the outcome	Hospital admission for an RTI episode proved to be laboratory-confirmed RSV	Hospital admission for an RTI episode proved to be laboratory confirmed RSV	Hospital admission for an RTI episode proved by laboratory confirmed RSV	Hospital admission for an RTI episode proved to be RSV by confirmed laboratory test
Results	- Infants admitted to hospital for an RSV proven-LRTI (n=66) -Infants with no LRTI hospital admission (n=1,692)	-Cases (n=186): infants admitted to hospital for an RSV proven-LRTI -Controls (n=371): infants admitted to hospital NOT related to RSV LRTI	- Infants hospitalized to hospital for an RSV proven-LRTI (n=202) - Infants with no LRTI hospital admission (n=5,239)	-Total Cases (N=145) -Total Controls (N=292) **33–35 wGA:** -Cases n=8 (5.6%): children admitted to hospital for an RSV proven-LRTI -Controls n=14 (4.9%): children with LRTI not due to RSV and not requiring hospitalization
RSV season analyzed	Two (2001–2003)	One (2002–2003)	Two (2005-2007)	Four (2000–2004)
Statistical analysis	Logistic regression analysis of all independent variables with p <0.15 in the univariate analysis	Logistic regression analysis of all independent variables with p <0.1 in the univariate analysis	Logistic regression analysis of all independent variables p <0.1 in any of the bivariate analyses was included in the multivariate model	Logistic regression analysis of all variables resulted to be significant in the univariate analysis or considered a priori important for the outcome

Note: Pediatric Investigators Collaborative Network on Infections in Canada **(PICNIC).**
-Risk Factors that Increase the Likelihood of Hospital Admission Caused by RSV Infection in Premature infants **(FLIP).**
-Osservatorio Virus Respratorio Sinciziale **(VRS).**

Table 2-3: Independent Risk Factor Related to RSV-LRTI Hospitalization Identified by Multivariate Logistic Regression Analysis in Different Studies

Risk Factor	PICNIC[11]	FLIP[20]	FLIP-2[21]
Chronological age	Birth in the first half of the RSV epidemic season (November– January) **OR (95% CI): 4.88 (2.57–9.29)**	Birth ±10 weeks the start of the RSV season (born between July 15 and December 15) **OR (95% CI): 3.95 (2.65–5.90)**	Absolute chronologic age at the start of RSV season ≤10 week **OR (95% CI): 2.99 (2.23– 4.01)**
Birth weight	Small for gestational age (birth weight <10th percentile) **OR (95% CI): 2.19 (1.14–4.22)**	**This variable was not considered a risk factor for this population**	**N/A**
Gender	Male gender **OR (95% CI): 1.91 (1.10–3.31)**	Male gender **Not significant at the bivariate analysis**	**N/A**
Breastfeeding	Not currently fed breast milk **This variable was not considered a risk factor for this population**	Breast-feeding duration of ≤2 months **OR (95% CI): 3.26 (1.96–5.42)**	**Not significant**
Older siblings	Any siblings <6 yrs **OR (95% CI): 2.76 (1.51–5.03)**	Any siblings >3 yrs **OR (95% CI): 2.85 (1.88–4.33)**	School age siblings or day care attendance++ **OR (95% CI): 2.04 (1.53–2.74)**
Day care attendance	Day care attendance **OR (95% CI): 12.32 (2.56-59.34)**	**Day care attendance not valuable because of too few cases**	**See above++**
Family history of allergies/ respiratory disease	Family history of eczema in the 1st degree relative **OR (95% CI): 0.42 (0.18–0.99)**	Family history of wheezing **OR (95% CI): 1.90 (1.19–3.01)**	**Not significant**
Crowding	>5 inhabitants (including the studied subject) in household during the RSV season **OR (95% CI): 1.69 (0.93–3.10)**	≥4 inhabitants and/or visitors (excluding siblings and case infant) in household during the RSV season **OR (95% CI): 1.91 (1.19–3.07)**	≥ 4 adults at home **Not significant**
Exposure to tobacco smoke	Exposure to tobacco smoke >1 smoker (any and total number of smokers in household) **OR (95% CI): 1.71 (0.97–3.00)**	At least one smoker (family member or person who usually cared for the child) **This variable was not considered a risk factor for this population**	Smoking during pregnancy **OR (95% CI): 1.61 (1.16 –2.25)**

Note: OR=odds ratio, CI=confidence interval.
-Osservatorio VRS as limited number of data available for 33–35 wGA infants specifically.

Table 2-4: Summary of Significant Risk Factors for RSV-Associated Severe LRTI Among Infants Born 33-35 wGA

Study	Significant Risk Factors
PICNIC	-Child care attendance -November–January birth -Preschool-aged siblings -Birth-weight <10% -Male gender -≥2 smokers in house ->5 people in household -Lack of eczema in first degree relative
FLIP	-Chronological age (CA) ≤10 week -Breast feeding ≤ 2 month -≥1 school-aged sibling -Family history of wheezing -≥4 adults in the home
FLIP-2	-Chronological age (CA) ≤10 week -Siblings or child care attendance -Tobacco exposure during pregnancy

Table 2-5: Canadian RSV Risk Scoring Tool for Infants Born Between 33 and 35 wGA:

	Score	
Risk Factor	*Yes*	*No*
1-Birth month is November, December or January (RSV season)	25	0
2-Subject or sibling attending daycare	17	0
3-More than five individuals in the home (crowding)	13	0
4-Small for gestational age (birth weight is less than the 10th percentile for gestational age)	12	0
5-Immediate family (mother, father, sibling) history without eczema (without eczema = yes)	12	0
6-Male sex	11	0
7-Two or more smokers in the household	10	0
TOTAL SCORE	100	-

CHAPTER 3: OBJECTIVES

1) To identify risk factors present at birth for RSV infection associated with hospital admission within one year of age among Nova Scotian infants 32 (+ 0 days) to 35 (+ 6 days) completed wGA.

2) To develop a model to predict hospitalization due to RSV of infants born at 32 (+ 0 days) to 35 (+ 6 days) wGA in Nova Scotia based on information present at birth.

CHAPTER 4: METHODOLOGY

4.1 OVERVIEW OF STUDY DESIGN AND ETHICS

The purpose of this retrospective cohort study was to identify risk factors present at birth in moderately premature infants 32 (+ 0 days) to 35 (+6 days) completed wGA in NS, Canada, for RSV-H. Using the unique provincial health card number (HCN) for each subject, the study used data obtained via a record linkage between a provincial perinatal database, a cardiology database, a database of Pz recipients, a National hospital admissions database and the Canadian Census database.

The cohort of infants were followed from birth to one year of age over 10 RSV seasons (1998 to 2008) allowing all children exposure to an RSV season. Risk factors were compared between infants hospitalized with an associated RSV diagnosis to the remaining infants with no RSV-H. Predictive variables were identified and combined into a predictive tool in order to identify infants at highest risk of hospitalization.

This thesis is a subset of a larger project evaluating the accuracy of an RSV-hospitalization prediction tool. That study also determined the number of Nova Scotian infants >32 weeks GA who would potentially be eligible for provincial government-funded receipt of Pz if an annual seasonal prophylaxis program were available.

The project received approval from the Joint Data Access Committee of the Reproductive Care Program (RCP) and received ethics approval from the Research Ethics Boards of the IWK Health Centre (project approval #1011135; approval date April 24th, 2012).

4.1.1 Inclusion and Exclusion Criteria

All NS born infants with a live birth between July 1st, 1998 and June 31st, 2008 whose gestational age was 32 weeks 0 days to 35 weeks 6 days gestational were included.

Infants were excluded if the birth was not in a hospital setting (i.e. since the information is not captured in a hospital database) and if the infant was previously administered RSV immunoprophylaxis (Pz). Infants with additional health conditions including CHD and patients with known CLD/BPD or were also excluded. In addition, infants who experienced early neonatal deaths (0-7 days) as well as late neonatal deaths (7-28 days) were excluded from the study. Any death <28 days would likely prevent the infant's exposure to community acquired RSV infection.

4.1.2 Defining Gestational Age

Gestational age is the time elapsed between the first day of a mother's last menstrual period (LMP) and the day of delivery. GA is usually measured in weeks (wGA). LMP is used to determine GA but other data can be used if this is not sufficient such as pregnancy tests, LMP, cycle length/regularity, contraception history, ultrasound information and assisted reproductive technology. LMP is felt to be an accurate method of dating the pregnancy if the patient is certain about the dates. Ultrasound dating is only used if there is an uncertain LMP or if there is discordance between menstrual and ultrasound assessment [71]. All data on GA is extracted from patient medical charts and prenatal forms by the province's RCP trained coders and entered accordingly into the Nova Scotia Atlee Perinatal Database (NSAPD) [72]. Infants were included for analysis as moderately premature if their gestation lied between 224 and 251 days (between 32 to 35 completed weeks GA).

4.1.3 Study Outcome-RSV-Associated Hospitalization

The purpose of the study was to find variables present in the neonatal period that predict hospitalization due to RSV in the first year of life. Although an unlikely possibility, a second RSV-H may occur in the same year. In order to keep analysis on variables predicting the first episode of RSV-H as our study endpoint, a second admission for RSV-H was not considered for analysis. Using the unique provincial HCN for each

subject, RSV-H and date of admission were determined through linkages to the NSAPD and the Canadian Institute for Health Information database (CIHI), respectively. RSV-H was determined using hospitalization discharge coding performed by trained individuals in accordance with the International Classification of Diseases, Ninth and Tenth Revisions (ICD-9-CM and ICD-10-CM). Analysis was restricted to only the primary outcome ICD-CM code diagnosis made by the attending physician or trained healthcare professional as the outcome measurement. Table 4-1 provides the validated ICD-CM codes that were used to identify RSV-H.

Narrowing the definition for RSV by selecting only RSV-specific codes was elected as the most conservative outcome definition; however, reliance on RSV-specific ICD-CM diagnoses is subject to an unknown magnitude of error. True cases of RSV may not have been captured because of failure to conduct a laboratory test for a respiratory tract infection proved to be RSV by viral culture and/or rapid test. It is possible that true RSV infected patients were coded as unspecified bronchiolitis or pneumonia, making the actual number of RSV cases even greater. This diagnostic uncertainty limits the validity of the definitions used for the outcome measurement, RSV-H. Although not as specific as hospitalization with laboratory confirmation of RSV, recent studies have used this definition for RSV-H [73, 74].

4.1.4 Predictive Variables

The literature presents a number of characteristics which are associated with increased risk for RSV-H (Section 2.4). Of these, 13 had similar or the same, variables in the Atlee. In addition to previous risk factors identified in the literature, Apgar score at both 1 and 5 minutes, mode of delivery, premature preterm rupture of membranes (PPRoM) and antenatal steroids were additionally examined; providing a total of 17 risk factors for analysis. To the author's knowledge these variables have yet to be investigated in association to severe RSV infection leading to hospitalization, and will add to the existing literature. Of the variables selected for analysis, some were identical to variables identified within the Atlee database, while others were modified to proxy variables.

Below is a detailed description of each risk factor identified in the literature. Based on this review, the final variables selected for analysis in this study are shown in Table 4-2.

Predictive variables identified in the literature and available in the Atlee

Birth month: Birth shortly before or early after onset of RSV season (In Canada RSV season usually begins between November and January and persists for four to five months) results in a longer exposure period to RSV and subsequently increased risk for infection [75]. The infant's birth month was used as a risk factor for acquisition of severe RSV leading to hospitalization.

Gestational age (GA): Both the AAP and CPS guideline recommend that RSV prophylaxis be considered for infants born at 32-35 weeks of gestation. The infant's GA measured in both weeks and days is a direct measurement collected from the Atlee determined as one the following:

-32 weeks+0 days-32 weeks+6 days (224 - < 231 days);
-33 weeks+0 days-33 weeks+6 days (231 - < 238 days);
-34 weeks+0 days-34 weeks+6 days (238 - < 245 days);
-35 weeks+0 days-35 weeks+6 days (245 - < 252 days);

Chronological age (CA): CA of the infant at the beginning of the RSV season is a known risk factor for RSV-H. Several factors play an important role with young CA, including the immature immune system and narrow airways [76]. The infant's CA at the beginning of highest RSV activity (determined as December 1st in NS, see results Table 5-3), was determined by using the subject's birth date (DD/MM/YYYY) and calculated from December 1st of the one year follow up.

Preschool aged siblings: The AAP recognizes the presence of school aged siblings (≤ 5 years of age) as a risk factor for RSV infection [70]. The underlying mechanism may be the increased contact intensity between siblings, resulting in the higher probability of transmission taking place due to prolonged exposure and closer proximity [75, 77]. The Atlee data does not collect the subject's number of preschool aged siblings; however the

number of pre-school aged siblings (≤ 5 years of age) will be measured by the mother's previous birth record (Gravida) and calculate births from previous deliveries in proceeding 5 years as a *proxy* for this risk factor.

Birth weight: Low birth weight has been found to add to the discriminatory power in identifying children susceptible for RSV hospitalization [78]. Birth weight was reported in grams (g) and included for risk factor analysis.

Small for gestational age: The role of intrauterine growth restriction has only been examined in the Canadian PICNIC study [11]. Category for birth weight for GA according to Kramer's criteria for Canadian premature infant's based on GA and sex of the infant (<3rd -97th percentile)) was used as direct measure [79]

Infant sex: It has been postulated that male infants have shorter and narrower airways than females infants, and are more likely to develop bronchial obstruction upon RSV infection, therefore requiring hospitalization [80]. Phenotypic sex (male or female) was used as direct measures from the Atlee.

Passive household smoke exposure: A recent review concluded that environmental smoke exposure increases the risk of hospitalization for RSV-attributable severe infection in infants [69]. Mother's smoking status at the time of delivery was used as a measurement for passive household smoke exposure. This also included the number of cigarettes smoked per day at the time of prenatal visit from 18-22 weeks as an additional analysis.

Household crowding: In a recent systematic review residential crowding significantly increased the odds of RSV hospitalization [81]. Development of RSV infection may be heightened within a crowded household environment. This would include an increased likelihood of exposure to infectious secretions after contacting individuals or surfaces contaminated with RSV. Household crowding (≥5 individuals per household including the study subject) was inferred from the number of previous deliveries in proceeding 18 years (gravida) and marital status (single vs. married) as a *proxy*.

Socioeconomic status (SES): Less parental education and low SES have also been reported to be a risk factor for RSV infection [32]. Mother's postal code linked to Canadian census data was used as an approximation for neighbourhood income quintile as a *proxy* for family socioeconomic status. This was based on postal codes converted to enumeration area with Statistics Canada estimates of neighbourhood income applied to each enumeration area. Categorization of mother's neighbourhood income quintile in the Atlee was performed by the RCP of NS.

Breast feeding: The protective role of breast feeding in preventing RSV-H is attributed to RSV-IgA and lactoferin[49]. Breast feeding was coded as infant breastfeeding at time of discharge from hospital (Yes/No). Breastfeeding was coded yes if the infant was breastfeeding and/or was being supplemented with formula at discharge.

Apgar scores at both 1 and 5 minutes: The Apgar rating is based on a total score of 1 to 10. A score of 7, 8, or 9 is normal and is a sign that the newborn is in good health while any score lower than 7 is a sign that the baby needs medical attention [82]. To the author's knowledge Apgar score as a risk factor for RSV-H has not yet been studied. Apgar scoring for newborn delivery between 0-10 at both 1 and 5 minutes was collected from the Atlee.

Mode of delivery: Regarding the route of the infant's birth, the level of definition for this project is "mode of delivery" which included a set of separate codes for either vaginal or caesarean section delivery. To the author's knowledge the association between vaginal and C-section delivery and poor RSV outcome has not yet been determined.

Maternal age: When compared with term infants, preterm infants with RSV infection were more likely to have an older mother in a recent study by Gooch et al [24]. The infant's mother age at the time of delivery was used as a direct measure.

Multiple births: Twins and triplets were reported as having a significantly higher risk of severe RSV infection [34]. The subject's birth type included as fetus number, singleton or multiple (combining twins and triplets), as a direct measurement from the Atlee.

Premature preterm rupture of membranes (PPRoM): To the author's knowledge PPRoM as a risk factor for RSV-H has not yet been analyzed in the literature. Information on PPRoM was collected from the Atlee as a direct source of measurement whether the subject's mother had or had not had not experienced PPRoM before the infant's birth.

Antenatal steroids: To the author's knowledge this will be the first study examining the role of antenatal steroid as a risk for severe RSV infection. Collected from the Atlee, the use of antenatal steroid during pregnancy was used as a source of direct measurement.

Predictive variables identified in the literature not available in the Atlee

Daycare attendance: Attendance of infants and young children in a child care setting or a day care group is regarded as a significant risk factor for RSV-H[11, 22, 23, 56, 83]. In the Canadian PICNIC study, day care attendance of the child was the single greatest risk factor in preterm infants born at a gestational age of 33-35 weeks. Additionally, the AAP recognizes childcare attendance and sibling(s) <5 years of age as a risk factor for RSV infection [70].

Family history of allergies/eczema/respiratory disease: In the Canadian PICNIC study, a history of eczema in a first degree family member was found to be an independent protective factor for RSV hospitalization in infants 33-36 wGA. The history of wheezing in the family was found to be statistically significant in the original The FLIP study, while the interaction between history of eczema in the family did not reach statistical significance in the multivariate logistic regression analysis. The FLIP-2 study concluded that family history of wheezing did not show statistical significance.

Low maternal RSV antibody levels: In preterm infants, incomplete maternally antibody transfer has been implicated to increase the risk for severe RSV infection. High titers of maternally derived RSV neutralizing antibodies have been suggested to be inversely associated with the incidence of severe RSV infection [23, 32, 84].

Race/ethnicity: Concurrent studies in American Indian/Alaskan Native and Inuit infants and children have shown RSV-H hospitalization rates that are almost 2 to 8 times higher in some of these populations, which may be attributable to isolation in northern or rural remote communities [30, 31].

There is evidence that daycare attendance, family history of allergies/respiratory disease, low maternal RSV antibody levels and race are associated with severe RSV infection and were therefore considered for analysis. However, the Atlee database does not include sufficient information on these potentially important risk factors.

Appendix A lists the risk factors discussed in the literature, and all the variables present in the Atlee database selected either direct measurement or proxy predictive variables for the statistical analysis.

4.2 DATA SOURCES

This study uses data from the (1) Nova Scotia Atlee Perinatal Database, (2) the IWK Cardiology Database, (3) the Nova Scotia Blood Coordinating Program database of Pz recipients, (4) the Canadian Institute for Health Information Hospital Admissions Database (CIHI), (5) and data from the Canadian census.

4.2.1 Reproductive Care Program (RCP)-The Nova Scotia Atlee Perinatal Database (NSAPD)

The Atlee began collecting data in 1988. It includes information on all in-hospital deliveries in NS as well as a small number of NS residents who deliver in New Brunswick. The RCP maintains the Atlee database, which is located at the IWK Health Centre in Halifax. The database contains extensive information on maternal, prenatal, labour, delivery and in hospital post-partum period. This includes information (mother and infant) on all deliveries of a live born or stillbirth infant weighing 500 grams or more [85]. As well, this database contains information on a number of variables that potentially increase the risk for RSV-H. The database also provides information on presence of co-morbidity that would confer eligibility for Pz, such as CLD of prematurity/BPD, and thus exclude an infant from this study.

Trained coders abstract and enter information into the NSAPD from standardized clinical prenatal forms, labour and delivery records, and post-partum records following established coding manuals. The extent of missing data for labour, delivery, and infant variables is close to 0%, while data on smoking is missing in about 5% of records. Data entry quality is regularly verified [85].

Data for linking the NSAPD included date of birth and provincial HCN. In order to construct the study cohort, date of birth (between June 1st, 1998 and July 31st, 2008), GA (clinical estimate of GA between 32 (+ 0 days) to 35 (+6 days)), and mother's postal code (as a measurement infant's residency in NS) were used to identify the cohort.

4.2.2 IWK Cardiology Database

The IWK Cardiology Database, housed and managed by the Division of Pediatric Cardiology the IWK Health Centre, contains data on all patients seen in the division extending back to 1966. The Cardiology database provided information on co-morbidity (CHD) that conferred eligibility for Pz, and the health records of all children with hemodynamically significant heart disease. These identified infants were removed from the study cohort. Data for linking the IWK Cardiology database included date of birth and provincial HCN.

4.2.3 Nova Scotia Blood Coordinating Program Database of Pz Recipients

The Provincial Blood Coordinating Program was created in January 2003. Initially developed at the NS Department of Health and Wellness, the database was transferred to the IWK in 2007 [15]. The database records all children who have received Pz in NS since and including 2003 and was used to determine which infants had received Pz prophylaxis, to which would make them ineligible. Data linking variables for the Provincial Blood Coordinating Program database of PZ recipient included date of birth, and provincial HCN.

4.2.4 Canadian Institute for Health Information Hospital Admissions Database (CIHI)

The Canadian Institute for Health Information database is comprised of data for each in-patient and out-patient hospital visit in NS from 1979 until present and provides variables necessary to estimate the outcome of interest, RSV-H. Data elements include HCN, date of birth, attending physician, hospital number, and any admission and discharge diagnosis classified by International Classification of Disease Clinical Modification code (ICD-9-CM and ICD-10-CM). The Canadian Classification of Health Interventions codes the main condition investigated or treated. Hospital admission date, hospital discharge date, date of service and Intensive Care Unit (ICU)/Neonatal Intensive Care Unit (NICU) admission are also available variables from CIHI [86, 87]. Data linking variables for the CIHI database included date of birth, and provincial HCN. Details on the outcome indicators used for analysis are discussed in 4.1.3.

4.2.5 Canadian Census Data

The Canadian census is conducted by Statistics Canada and takes place every five years [88]. Mother's postal code linked to Canadian census data was used as an approximation for neighbourhood income quintile. Categorization of mother's neighbourhood income quintile in the NSAPD was performed by the Reproductive Care Program of NS.

Appendix B provides a detailed description of all variables included for defining the study eligibility criteria, predictive variables for follow up analysis and outcome of interest

4.2.6 Data Linkage and Security Measures

The study cohort was previously assembled using the infant's provincial HCN, and infant's date of delivery, by an authorized researcher through the RCP who linked the dataset. The final cohort was constructed after the inclusion of several additional risk factors from the NSAPD. This was again done by an authorized researcher through the RCP. All major identifiers such as health card numbers, names, birthdates, and postal codes were stripped before the restricted access was given to the analysis file.

4.3 MINIMAL DETECTABLE ODDS RATIOS

Power analysis was used to calculate the minimum effect size of risk factors likely to be detected given the estimated sample size and expected hospitalization rate for the retrospective study. The study's population was pre-determined by the total number of live births for infants between 32 and 35 completed wGA in NS. It is expected that approximately 2,800 infants with about 90 RSV-H events will ensue over the ten years of the study (rate of RSV-H= 3.2%).

Analyses were performed assuming a level of significance (α) of 0.05 and a power (1-β) of 80%. The minimal detectable odd ratios (ORs) were calculated in both the univariate and the multivariable analysis (binomial and five ordinal predictors). Power analyses calculation were tested against a conservative 2-3.5% likelihood of being hospitalized given that expected rate of hospitalization was 3.2%. An OR of 1.5 indicated acceptable levels of power 0.845 and an OR 3 or greater was indicative of 100% power. Univariate and multivariate power analysis were conducted using models that included ordinal, categorical and continuous predictive variables. Based on these analyses, in order to calculate an OR of 1.5 for a categorical variable a power of 0.845 was acceptable. An OR of 1.5 significant at (α) of 0.05 for any binary categorical explanatory variable would have power of 0.845 to be detected. For any ordinal or continuous variable the power would be indicative of a higher level of power.

4.4 STATISTICAL ANALYSIS

All analyses were done using SAS for windows version 9.2. The following section outlines analysis for both research objectives. Unless noted otherwise, statistical significance for all analyses was defined at the 2-sided P=0.05 level.

4.4.1 Objective 1: Identify Risk Factors for RSV Infection Associated with Hospital Admission Within One Year of Age

Table 4-2 describes the treatment of the predictive variables included for analysis. Certain variables were transformed into categorical values. Chronological age was retained as continuous for the logistic regression analyses however to estimate the relative risk for CA in the bivariate analysis, CA was transformed into a categorical variable (based on the results from the Student's t-test). Table 4-3 describes the treatment of the outcome variable included for analysis. The outcome variable was dichotomized as either RSV-H or no RSV-H.

To identify predictive variables associated with RSV-H, Cochran-Mantel-Haenszel $\chi 2$ statistics and odds ratios using 95% confidence intervals (CI) were calculated for all predictive variables in the univariate logistic regression. To be conservative, all predictive variables with an unadjusted relative risk at the significant level $p<0.15$ were included in the initial multiple logistic regression model. Adjusted relative risks were determined through multivariate logistic regression. The logistic regression model used backward elimination as the effect-selection method. Final predictors were considered significant using the Likelihood Ratio (LR) test. Overall model evaluation was tested using the LR, Score, and Wald tests. The Homer & Lemeshow test was used to determine goodness-of-fit.

4.4.2 Testing and Interpreting Coefficients in the Multiple Regression Analysis Containing Interactions

To test the theoretical model explaining which predictive variables were significantly associated with RSV-H, emphasis was put on interpreting any interaction effects on the dependent variable, RSV-H. To test for two-way interactions (often thought

of as a relationship between an independent variable and dependent variable, moderated by a third variable (moderator)) maximum likelihood estimates based on the Fisher's scoring technique was performed in order the examine any significant interaction(s). The strongest interaction effects were forced into the model containing the statistically significant independent variables from section 4.4.2. A significant interaction was included only if its inclusion resulted in a significant change in the –2 Log Likelihood from the previous step. As such, the final model consisted of factors that were statistically significant or interactions that were forced into the model. This second model was compared to the first exploratory, statistically driven model.

4.4.3 Objective 2: Develop a Model to Predict RSV-H of Infants Born at 32-35 wGA in Nova Scotia

In order to develop a user-friendly clinical instrument that would predict the risk of RSV-H, the significant variables included in the final logistic regression model were used in developing a scoring system. Final predictors were considered significant using the LR test and conditional parameter estimates and RSV-H as the dependent dichotomous outcome variable.

By calculating the parameter estimates of the logistic model, each risk factor was assigned an estimated independent weight equal to its agreeing parameter estimate in the multivariable model ($f(Y) = \beta 0 + \beta 1\ X1 + \beta 2X2 + \ldots + \beta pXp$). Since the variables in our parameter regression analysis were measured in different units, standardization of the coefficient (β) was done so that the coefficients ignore the independent variable's scale of unit (example infant weight in grams and family crowding in number of individuals).

Each variable in the model was then transformed into a binary predictor classified as either absent (no) or present (yes) and placed on a dichotomous scale and converted into a positive value, based on the standardized β coefficient. To produce the predictive instrument, the independent standardized weights for each individual risk factor (β coefficient) were converted into a relative weight by calculating each individual risk factors β coefficient over the cumulative risk estimate in the cohort. The re-weighted standardized parameter (β coefficient) from the standardized logistic regression was

converted into variable scores (multiples of 100) and rounded off to the nearest integer. This provided each patient a score from 0-100 based on the cumulative risk of RSV-H [54].

4.4.4 Receiver Operator Characteristic (ROC) and Test Characteristics

Receiver Operator Characteristics and the Area Under the Curve

The Receiver Operator Characteristic (ROC) was used in finding the score that produced the highest accuracy in predicting cases of RSV-H (the highest ***sensitivity*** and lowest false positive ratio (1-***specificity***));

Sensitivity is the proportion of truly hospitalized infants who were correctly identified by the scoring tool, i.e. who have high scores.

Specificity is the proportion of infants without RSV-H correctly identified by the scoring tool, i.e. they have low scores.

The curve was obtained by calculating the sensitivity and specificity of the scoring tool at every possible cut-off point, and plotting sensitivity (y-axis) against 1-specificity (x-axis). For clinical practice, selecting the cut-off point(s) to discriminate between those with low risk (low scores) and those with higher risk (high scores) is optimal. One method commonly used assumes that the best cut-off point for balancing sensitivity and specificity of a test is the point on the curve closest to the (0, 1) point [89,90]. This method was applied in order to determine appropriate cut-off values for low risk to the highest at-risk groups. The information was used to produce the final scoring tool.

The area under the curve (AUC) serves as a single measure, independent of prevalence, and summarizes the discriminatory ability of the scoring tool across the full range of cut offs. A perfect test would have an AUC of 1.0, or 100% predictive probability, while a useless test (one whose curve falls on the diagonal line) has an AUC of 0.5 [91]. The AUC was interpreted as a reflection of how good the test distinguished between RSV-H and non RSV-H infants.

Estimating the Sensitivity and Specificity of the Scoring Tool

The accuracy of our scoring tool was characterized by its sensitivity and specificity for predicting either an event (RSV-H) or a non-event (no RSV-H). Results were tabulated using a 2 x 2 table (See Results Table 5-12 and 5-13). The table classifies infants who were either true positive (TP = # of correctly predicted events); true negative (TN = # of correctly predicted non-events); false positives (FP = # of non-events predicted as events); and false negatives (FN = # of events predicted as non-events). From these we calculated the sensitivity and specificity;

- **Sensitivity = TP ratio = TP / (TP+FN)**
- **Specificity = TN ratio = TN / (FP + TN)**

Positive and Negative Predictive Values

The scoring tool was further assessed by calculating the probability that an assigned score would give the correct diagnosis. Positive and negative predictive values (PPV and NPV) describe a patient's probability of having either a positive or negative outcome once the result of the scoring tool is assigned.

PPV is the overall probability of an infant having RSV-H when the scoring tool is positive; PPV is equal to the TP divided by the sum of the TP and FP;

- **PPV = TP / (TP+FP)**

NPV is the overall probability that an infant will not have an RSV-H when the scoring tool is negative; PPV is equal to the TN divided by the sum of the TN and FN;

- **NPV = TN / (TN+FN)**

Likelihood Ratio

The estimated probability of disease *before* the specified test result is known is referred to as the **pre-test probability.** The patient's probability or chance of having the disease *after* the test results are known is referred to as the **post-test probability** [92]. The

sensitivity and specificity of the scoring tool were combined into a measure called the likelihood ratio (LR). The LR was calculated in order to determine whether knowing the results of the scoring tool (post-test probability) would improve on a clinical guess based simply on pre-test probability (i.e. how much does the probability of having the disease increase following a high score?). Two versions of the likelihood ratio exist, one for positive and one for negative test results. They are known as the *positive likelihood ratio* (LR+) and *negative likelihood ratio* (LR–), respectively. Likelihood ratios for both a positive test (LR+) and a negative test (LR–) were calculated by using the sensitivity and specificity;

- **LR + = TP rate / FP rate**
  = Sensitivity / (1-Specificity)
- **LR – = FN rate / TN rate**
  = (1-Specificty) / Sensitivity

If the LR+ or LR – is equal to 1, it would imply that the pre-test and post-test are similar, and the therefore the results add no additional information. Conversely, a LR>1 indicates that the post-test probability is greater than the pre-test probability. Likelihood ratios above 10 and below 0.1 would reflect large differences from the pre-test probability [93, 94].

Estimating Probability of Disease: Post-Test Probability

Sensitivity, specificity and LRs, were used to calculate alternative statistics for summarizing diagnostic accuracy of the scoring tool. Bayes theorem was applied in order to obtain the post-test of RSV-H by multiplying the pre-test odds by the LR of the test. [93]

- **Post-test odds = pre-test odds x likelihood ratio**

By assuming the pre-test probability of 3.1% (88 RSV-H / 2,811 non-RSV-H), the post-test probability for RSV-H was calculated. The use of the odds rather than probabilities was calculated as follows (the pre-test probabilities were converted to pre-test odds then multiplied by LR to get the post-test odds, which was then converted into post-test probabilities);

- **Pre-test probability = ρ_1 = 0.03 (~3% probability of RSV-H)**
    - **pre-test odds = $\rho_1/(1- \rho_1)$**
    - **Post-test odds = o_2 = pre-test odds x likelihood ratio**
    - ***Post-test probability = $o_2/(1+ o_2)$***

The results from the ROC, AUC, LR+/-, PPV, NPV and post-test probability were used as additional statistics in order to interpret the results and assess the value of the RSV-H scoring tool for clinical practise.

Table 4-1: Treatment and Frequency of Outcome Variables in the CIHI

Outcome Measurement	ICD-CM Code	Description on ICD-CM Outcome Variable Measurement
First hospitalization associated with RSV diagnosis during the one year follow up	**ICD-9-CM***	
	079.6	RSV infection
	466.11	Acute bronchiolitis due to RSV
	480.1	Pneumonia due to RSV
	ICD-10-CM**	
	B97.4	RSV as the cause of diseases classified to other chapters
	J12.1	RSV pneumonia
	J21.0	Acute bronchiolitis due to RSV

* The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) [95].
** The International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) [96].

Table 4-2: Treatment of Predictive Variables Extracted from the Atlee Database

Predictive Variable (Literature)		Variable Available from the Atlee		Direct measure or proxy	Treatment of Atlee Variable in this Study
Variable	***Description***	***Variable***	***Description***		
Month of birth	-	Date of birth	Month of birth identified from the infant's DOB (MM/DD/YYYY)	Direct measure	Retained as twelve categories: September, October, November, December, January, February, March, April, May, June, July, August
Gestational age (GA)	32 weeks + 0 days -35 +6 days weeks GA	GA of infant at birth in weeks/days	-32 weeks: (224 - < 231 days) -33 weeks: (231 - < 238 days) -34 weeks: (238 - < 245 days) -35 weeks: (245 - < 252 days)	Direct measure	Collapsed into four categories: 32 weeks, 33 weeks, 34 weeks and 35 weeks
Chronological age (CA)	CA at the start of the RSV season	Date of birth (MM/DD/YYYY)	DOB used as a calculation of subject's CA in weeks on December 1st during the one year follow up after birth.	Direct measure	Maintained as continuous variable
Preschool aged siblings	Infant's siblings that are ≤ 5 years of age	Gravida of the subjects mother	Number of previous births of children ≤ 5 years of age and the assumption made that they are still in the household	Proxy	Dichotomized: no siblings ≤5 years of age and ≥ 1 siblings ≤ 5 years of age
Birth weight (grams)	-	Birth weight in grams (g)	-	Direct measure	-Dichotomized: Normal birth weight (>2500g) and low birth weight infant (≤ 2500g)
Small for gestational age (percentile)	Category for birth weight for GA	Weight for gestational age (percentile)	Kramer's criteria for Canadian premature infant's based on GA and sex of the infant percentile (<3rd %-97th %) [79]	Direct measure	Dichotomized: ≤ 10^{th} percentile and >10^{th} percentile
Infant sex	Male or female	Phenotypic sex	Male or female	Direct measure	Dichotomized: male and female
Environmental tobacco smoke exposure	Passive household smoke exposure	Mother's smoking status at delivery date	Used as a measurement for passive household smoke exposure. Also included was the number of cigarettes smoked per day at the time of prenatal visit	Direct measure	Dichotomized: non-smoker and smoker (including 0.5 pack of cigarettes/day, < 1 but > 0.5 pack, ≥1 pack, and amount unknown)

Predictive Variable (Literature)		Variable Available from the Atlee		Direct measure or proxy	Treatment of Atlee Variable in this Study
Variable	***Description***	***Variable***	***Description***		
Household crowding/ number of siblings	Household crowding: ≥5 individuals per household including study subject	Household crowding: ≥5 individuals per household including study subject	Inferred from the number of previous deliveries in proceeding 18 years (gravida) and marital status (single vs. married) at the time the baby was born	Proxy	-**Marital status:** Dichotomized: no partner (including single, widowed, divorced, or separated) and partner (including married or common-law) -**Gravida:** Collapsed into total number of siblings ≤18; 0 siblings, 1 siblings, 2 siblings, 3 siblings, 4 siblings, 5 siblings, 6 siblings and 7 siblings
Low socioeconomic status (SES)/ parental education	-	SES	Mother's postal code linked to Canadian census used as an approximation for neighbourhood income quintile (performed by the RCP)	Proxy	Neighbourhood income collapsed into quintiles: highest, upper-middle, middle, lower-middle and lower
Breast feeding	-	Mother's breast feeding status at birth	Mother's Breastfeeding status indicated by the mother at the time of birth admission	Direct measure	Dichotomized: Breastfeeding (combining exclusive breastfeeding (E), combined breast milk (S) or other supplements and not breast fed (N)) and no breastfeeding.
Maternal age	-	Maternal age	Mother's age at time of delivery	Direct measure	Dichotomized: ≤25 years old and >25 years old
Multiple births	Singleton, twins or triplets	Number of fetuses	Subject's birth type; singleton, twins or triplets	Direct measure	Dichotomized: Singleton vs. multiple (twin and triplets combined)
NA	NA	Apgar score at 1 and 5 minutes	Score between 0-10 both at 1 and 5 min	Direct measure	Dichotomized: Normal score (7-10) and low score (0-6) both at 1 and 5 min
NA	NA	Mode of delivery	C-section or Vaginal	Direct measure	none
NA	NA	Premature preterm rupture of membranes (PPRoM)	PPRoM during current pregnancy	Direct measure	Yes or no
NA	NA	Antenatal steroids	The use of antenatal steroid during pregnancy	Direct measure or proxy	Yes or no

Table 4-3: Treatment of Outcome Variable in the CIHI

Variable	Database	Description	Treatment of Variables in the CIHI
RSV-associated hospitalization (RSV-H)*	CIHI	Infant's first hospitalization attributable to RSV-associated severe LRTI	Dichotomized: no RSV-H and RSV-H: (including ICD-9-CM codes 007.9, 466.1, 480.1 and ICD-10-CM codes B97.4, J12.1, J21.0)

***Note: RSV-H** = RSV-associated hospitalization, the dependent bivariate outcome measurement.

CHAPTER 5: RESULTS

5.1 DESCRIPTION AND CHARACTERISTICS OF STUDY COHORT

Of the 2,909 infants with a live birth in NS between July 1, 1998 and June 31, 2008 and GA between 32 weeks (+ 0 days) to 35 weeks (+6 days), there were 85 infants excluded because they had been vaccinated. Within the 85 vaccinated infants who were excluded, one infant had a cardiac indication, and 10 had documented lung disease. Lastly, a total of 10 infants were excluded for early neonatal death (0-7 days) and 3 for late neonatal death (7-28 days). The final eligible study cohort consisted of 2,811 infants.

Table 5-1 provides a detailed description of the characteristics of the study cohort. Missing values were identified in seven of the 17 variables. Mother's smoking status at admission was missing in 6.1% of records, and marital status in 3.5%. Breast feeding, family SES, Apgar scores and birth weight for GA were missing 2% of records. The mean (SD, range) for birth weight was 2,366 grams (514; 256-4695) and the mean (SD, range) of mother's age was 29 years old (5.7; 14-49).

5.1.1 RSV-Associated Hospitalization in the First Year of Life

During the first year of life, a total of 88 infants had a hospital admission associated with RSV- H (rate of 3.1%). During the study's ten RSV seasons, the highest number of hospitalizations in any one year occurred during 2006 (n=22). Most of the RSV-H diagnosis were coded as acute bronchiolitis due to RSV with 64 out of the 88 (73%) infants hospitalized were identified by ICD-10-CM J21.0. The mean (SD, range) chronological age on December 1st for RSV-H infants was 31 weeks (16.1; 0-52). Table 5-2 provides a detailed description of the outcome variable for infants who experienced a hospitalization associated with RSV. Results indicate that infants who were born in December, January and February experienced the highest rates of RSV-H (infants hospitalized with RSV/total number of infants born in the same month) with 4.9% (11/227), 6.5% (15/230) and 6.5% (13/199) hospitalization, respectively (Table 5-3).

5.2 DATA ANALYSIS-OBJECTIVE 1: IDENTIFY RISK FACTORS FOR RSV INFECTION ASSOCIATED WITH HOSPITAL ADMISSION WITHIN ONE YEAR OF AGE

5.2.1 Bivariate Analyses of the Predictive Variables

Table 5-4 shows the incidence of RSV-H and the relative risk (RR) estimates for all categorical predictive variables tested against the binary outcome as well as results from the Student's t-test for CA.

Nine predictive variables were found associated with RSV-H in the bivariate analysis at ($p<0.15$), and therefore eligible for inclusion in the multivariate logistic regression model. Infants born in December, January and February, ≥ 1 preschool age sibling in the household, low birth weight (≤ 2500g), passive household smoke exposure, household crowding (≥ 5 individuals in the household), no breastfeeding, no PPRoM, and CA (>31 weeks on December 1st) were all found to increase the likelihood of RSV-H. Lower-middle SES demonstrated a protective effect against hospitalization.

5.2.2 Logistic Regression Analyses of the Predictive Variables

All nine significant variables in the bivariate analysis were included as covariates in the backward stepwise multivariate logistic regression model. A total of 328 infants were deleted due to missing values for the risk factors included in the logistic regression analysis. The logistic procedure included 76 infants with RSV-H and 2,407 infants with no RSV-H.

After adjustments in the multivariate regression model, birth during December, January or February, passive household smoke exposure and household crowding remained independently associated with RSV-H (Table 5-5).

Interaction Effects Between Predictive Variables in Association with RSV-H

Three of the variables found to be predictive were also found to interact (Table 5-6 A-C). As seen in Table 5-6A, the strongest interaction effect was found between passive smoking and infants born in December-February ($p<0.0001$). The two-way

interaction indicates a higher hospitalization rate for infants born in Dec.-Feb. (6.1%) compared to those born Mar.-Nov. (1.5%) when there was no passive household smoke exposure. When smoking was present, infants born Dec.-Feb. experienced the highest rates at 6.3% while the hospitalization rate for infants born Mar.-Nov. remained high at 5.0 %. This is indicative of a multiplicative interaction effect between passive household smoke exposure and infants born in December, January and February.

Infants with low Apgar score at 1 min and who experience passive household smoke exposure together have the highest likelihood of RSV-H (6.1%) as demonstrated in Table 5-6B. Although infants with higher Apgar scores who have exposure to smoking experience slightly lower rates, these infants still experience overall high rates of hospitalization (5.2%). In comparison, infants with both low and high Apgar scores who do ***not*** have exposure to smoking experienced 3.5% and 2.4% rates of RSV-H respectively (p=0.0065).

Infants who were small for GA (≤ 10th percentile) that also experienced passive household smoke exposure had the highest rates of RSV-H (6.3%) as seen in Table 5-6C. This rate remained relatively high in infants with higher birth weights and exposure to smoking (5.2%). Infants with both low and high weight for gestational age who were *not* exposed to smoking did much better (3.1% and 2.5%) respectively (p=0.0094).

5.2.3 Testing the Significance of the Interactions in the Predictive Model

The three statistically significant interactions were considered when interpreting the results for the final predictive model. We ran regression models in which we entered all three significant interaction terms, as well as a combination of two of the three interactions and lastly models with only one interaction all included into the original three variable predictive model. Passive household smoke exposure and infant birth in Dec.-Feb. in relation to the RSV-H was found just shy of the level of significance, $p<0.05$, when included separately within the three variable model (p=0.0510).

Comparing the Model Fit Statistics of Both Models

The three variable model was compared to the three variable + interaction (passive household smoke exposure and infant birth in Dec.-Feb.) model. Both models were statistically comparable demonstrated by Somers' D statistic (0.369 and 0.370 respectively). To be applicable to clinical practice, a predictive model must discriminate between infants who will be hospitalized for RSV and those who will not be hospitalized. To keep to a practical score sheet for clinical practice, the three variable predictive model *without* the interaction term was used for the development of the RSV-H scoring tool. Tables 5-7 A and B provide full details on the model fit statistics for both models.

5.3 DATA ANALYSIS-OBJECTIVE 2: DEVELOP A MODEL TO PREDICT HOSPITALIZATION DUE TO RSV OF INFANTS BORN 32-35 WGA IN NOVA SCOTIA

5.3.1 Development of the RSV-H Scoring Tool

The calculation of the sum of the individual variable scores was discussed in section 4.4.3. The independent standardized weights for each individual risk factor (β coefficient) in the three variable predictive model were converted into relative weights. The re-weighted standardized parameters were converted into variable scores (multiples of 100) and rounded off to the nearest integer providing each patient a score from 0-100. Table 5-8 shows the final scores for each predictive variable in the RSV-H scoring tool.

5.4 RECEIVER OPERATOR CHARACTERISTIC (ROC): ANALYSIS FOR THE DEVELOPMENT OF THE RSV-H SCORING TOOL

There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool—household crowding, passive household smoke exposure and birth in December, January or February— therefore were excluded for the final RSV-H analysis. The remaining cohort included 78 infants with RSV-H and 2,478 infants with no RSV-H for the RSV-H scoring tool analysis.

Determining the Optimal Cut-Off Points and Assessing the Scoring Tools Predictive Accuracy

The ROC was used for two main purposes:

1- Provide determination of the cut-off point(s) at which optimal sensitivity and specificity were achieved.

2-Provide an assessment of the diagnostic accuracy of the overall scoring tool (AUC).

The ROC is shown in Figure 5-1 while Table 5-9 provides details on the ROC probability level, sensitivity, 1-specificity and the associated RSV-H scoring range (0-100). Results indicate that the highest sensitivity (true positive rate) and lowest false

positive rate were 0.705 and 0.397, respectively, corresponding to a logistic predicted probability estimate of 0.035 (~3.5%). The predicted probability identifies the RSV risk scoring cut-off value of 41, providing the highest accuracy in identifying between infants with relatively low risk for RSV-H compared to infants with a higher risk (≥ 42 cumulative score).

In order to further characterize the second group into moderate-risk and high-risk infants (RSV prophylaxis provided to the absolute highest risk infants would be most cost effective), a second cut-off among infants was attempted based on the scoring frequency. Highest sensitivity and lowest false positive rates were selected as 0.154 and 0.079, respectively, corresponding to a logistic regression predicted probability estimate of 0.0648 (~6.5%) which matches to a second cut off score of 64. As shown in Figure (Figure 5F-1) the AUC was 0.685. In general, the closer the AUC is to 1.0, the better the overall diagnostic performance of the test, and the closer it is to 5.0, the poorer the test [91].

The number and percentages of RSV-H for each risk group and their corresponding score cut-off are provided in Table 5-10. The scoring tool distinguished between low risk (infants with a 1.5% hospitalization rate) and higher risk infants (~5-6% hospitalization rate) based on the three significantly independent risk factors, however it fails to further distinguish between moderate and the highest at risk infants for RSV-H (5.2% and 5.8% probability of RSV-H, respectively). Lastly, Table 5-11 provides the mean predicted probabilities for the low, moderate and high risk groups, as well as the relative risk estimates for RSV-H categorization (low vs. moderate risk and low vs. high risk).

Assessing the Value of the Scoring Tool in Predicting Hospitalization

In order to answer the question "how much does the probability of having RSV-H increase once an infant is categorized in a risk group?" the LR+ was assessed. The TP, TN, FP and FN were used to calculate sensitivity, specificity. Sensitivity and 1-specificty were used to calculate the LR+ for low, moderate and high risk groups. An LR of 1 means that the post-test probability is the exact same as the pre-test probability (3.1%

RSV-H), while a LR less than 1 decreases the probability of disease. Table 5-12 indicate LR+ of 1.95, 1.73 and 0.49 for the high, moderate and low risk groups, respectively. The positive post-test probability of RSV-H was then calculated by multiplying the pre-test odds (0.032) of obtaining RSV-H and the respective LRs for each level of risk for the RSV-H scoring tool (Table 5-12). The LRs+ and the positive post-test probabilities between high and moderate risk groups indicate very similar increases in the probability of RSV-H, while the low risk group demonstrates a decrease in the probability of RSV-H.

Two 2x2 tables were constructed in order to further test the sensitivity and specificity of the 41 and 64 point cut-off. TP, TN, FP and FN were used to calculate PPV and NPV for both cut-off points. Results from the 2x2 table for the 41 point cut-off between moderate risk and high risk (>42) and low risk (0-41) indicate a 71% sensitivity and 60% specificity. The corresponding LRs are 1.78 and 0.56 (Tables 5-13). The 64 point cut-off between high risk (65-100) and moderate and low risk (42-64) provided 15% sensitivity and 92% specificity. The corresponding LRs are 1.88 and 0.95 (Table 5-14). A post-test probability of 5.5% was calculated between moderate and high and low risk groups, while the post-test probability between high risk and moderate and low risk was similar, at 5.8%.

In other words, the scoring tool was capable of identifying between low risk infants and moderate to high risk infants, however, the risk factors that make-up the RSV-H scoring fail to further identify infants at highest risk of RSV-H; those who would be the best candidates for Pz prophylaxis.

Table 5-1: Characteristic of the Study Cohort at Baseline

*Cohort Characteristics**	*N*	*%*	*Missing Values (%)*
Birth Month (RSV season)			
Born in December, January, February	652	23.0	0
Born March-November	2,159	77.0	
Gestational age (GA)			
32 weeks+0 days-32 weeks+6 days (224 to < 231 days)	257	9.1	0
33 weeks+0 days-33 weeks+6 days (231 to < 238 days)	422	15.0	
34 weeks+0 days-34 weeks+6 days (238 to < 245 days)	762	27.1	
35 weeks+0 days-35 weeks+6 days (245 to < 252 days)	1,370	48.7	
Chronological age (CA) at the start of RSV season**			
≤ 4 weeks (≤ 1 month of age)	207	7.4	0
> 4 to ≤12 weeks (>1 month to ≤ 3 months of age)	432	15.4	
> 12 to ≤ 24 weeks (>3 months to ≤ 6 months of age	674	24.0	
> 24 to 52 weeks (>6 months to one year of age)	1,498	53.2	
Gravida ≤ 5 years (Preschool aged siblings)			
0 siblings	1,509	53.7	0
1 sibling	921	32.8	
2 siblings	337	12.0	
3 siblings	44	1.6	
Birth weight (grams)			
Normal birth weight (> 2500g)	2,076	74.0	6 (0.2%)
Low birth weight (≤ 2500g)	729	26.0	
Birth weight for gestational age			
< 3rd percentile	87	3.1	6 (0.2%)
3rd to < 5th percentile	47	1.7	
5th to < 10th percentile	162	5.8	
10th to < 50th percentile	1,093	39.0	
50th to < 90th percentile	1,107	49.5	
90th to < 95th percentile	158	5.6	
95th to < 97th percentile	52	1.9	
> 97th percentile	99	3.5	
Infant sex			
Male	1,303	46.4	
Female	1,508	53.6	
Passive household smoke exposure (Mother's smoking status)			
Yes	643	24.3	161 (6.1%)
No	2,007	75.7	
Among smokers, pack (# cigarettes)/day at admission			
0.5 pack pack/day (1 - < 12 cigarettes)	307	47.7	-
< 1 but > 0.5 pack/day (12 < - < 25 cigarettes)	178	27.7	
≥1 pack/day (25 - 87 cigarettes)	49	7.6	
Amount unknown/day	109	17.0	
Marital status			
Married/Common-law	1,911	70.0	98 (3.6%)
Single/Divorced/Separated	802	30.0	

*Cohort Characteristics**	*N*	*%*	*Missing Values (%)*
Gravida ≤ 18 years			
0 siblings	1,284	45.7	0
1 sibling	883	31.4	
2 siblings	436	15.5	
3 siblings	137	4.9	
4 siblings	45	1.6	
5 siblings	19	0.7	
6 siblings	6	0.2	
7 siblings	1	0.04	
Household crowding (≥ 5 individuals per household)			
Yes	497	18.3	98 (3.6%)
No	2,216	81.7	
Socioeconomic status (SES)			
Highest	636	22.9	33 (1.2%)
Upper-middle	535	19.3	
Middle	499	18.0	
Lower-middle	572	2.6	
Lower	536	19.3	
Breast feeding			
Yes	1,765	63.8	45 (1.6)
No	1,001	36.2	
Apgar score at 1 minutes			
Normal score (7-10)	2,140	76.5	13 (0.5%)
Low score (0-6)	658	23.5	
Apgar score at 5 minutes			
Normal score (7-10)	2,702	96.6	15 (0.5%)
Low score (0-6)	94	3.4	
Mode of delivery			
Vaginal	1,649	58.7	0
Cesarean section	1,162	41.3	
Maternal age			
19 years or younger	123	4.4	0
20 to 24 years	492	17.5	
25 to 29 years	810	28.8	
30 to 34 years	885	31.5	
35 years or older	501	17.8	
Number of fetuses			
Singleton	2,130	75.8	0
Twin	623	22.2	
Triplet	58	2.1	
PPRoM			
Yes	992	35.3	0
No	1,816	64.7	
Antenatal steroids			
Yes	995	35.4	0
No	1,816	64.6	

*The final study cohort consisted of 2,811 infants.
**Subject's CA in weeks on December 1st during the one-year follow up after birth.

Table 5-2: Detailed Characteristic Description of Outcome Variables in the Study Cohort

*Cohort Characteristics**	*N*	*%*	*Missing Values (%)*
RSV-associated hospitalization per fiscal year (N=88)*			
1998	3	3.4	-
1999	4	4.6	
2000	9	10.2	
2001	-	-	
2002	11	12.5	
2003	13	14.8	
2004	17	19.3	
2005	6	6.8	
2006	22	25.0	
2007	1	1.1	
2008	2	2.3	
RSV-associated hospitalization (ICD-CM code) (N=88)*			
ICD-9-CM 007.9 (Viral infection)	1	1.1	-
ICD-9-CM 466.1 (Acute bronchiolitis)	13	14.8	
ICD-9-CM 480.1 (Pneumonia due to RSV)	2	2.3	
ICD-10-CM B97.4 (RSV)	6	6.8	
ICD-10-CM J12.1 (RSV pneumonia)	2	2.3	
ICD-10-CM J21.0 (Acute bronchiolitis due to RSV)	64	72.7	

* RSV-associated hospitalization (RSV-H) is the dependent bivariate outcome measurement. A total of 88 per 2811 (3.1%) infants in the study cohort were admitted to hospital associated to RSV during the study's 10-year period.

Table 5-3: Risk of RSV-Associated Hospitalization by Infant's Birth Month (1998-2008)

	Birth Month												
RSV-H*	**Oct.**	**Nov.**	**Dec.**	**Jan.**	**Feb.**	**Mar.**	**Apr.**	**May**	**Jun.**	**Jul.**	**Aug.**	**Sept.**	**N Total**
No (n)	224	216	215	212	186	243	223	242	243	253	240	226	2723
Yes (n)	7	5	**11**	**15**	**13**	8	5	3	3	5	8	5	88
Total N	232	222	227	230	199	251	230	245	246	260	260	232	2,811
% RSV-H	3.0	2.3	**4.9**	**6.5**	**6.5**	3.2	2.2	1.2	1.2	1.9	3.2	2.2	-

*A total of 88 infants with RSV-associated LRTI were admitted to hospital during the study's 10 year period.

Table 5-4: Incidence and Unadjusted Relative Risk Estimates of RSV-Associated Hospitalization

Variable	*% RSV-H*	*URR*	*95% CI*	*P-value*
Birth Month (RSV season) *				
Born in December, January, February	5.98 (39/652)	**2.64**	1.75-3.98	**0.0002**
Born March-November	2.27 (49/2,159)	1	-	-
Gestational age (GA)				
32 weeks+0 days-32 weeks+6 days (224 to < 231 days)	1.95 (5/257)	0.72	0.28-1.82	0.5876
33 weeks+0 days-33 weeks+6 days (231 to < 238 days)	4.50 (19/422)	1.66	0.97-2.87	0.1345
34 weeks+0 days-34 weeks+6 days (238 to < 245 days)	3.54 (27/762)	1.31	0.80-2.14	0.3528
35 weeks+0 days-35 weeks+6 days (245 to < 252 days)	2.7 (37/1,370)	1	-	-
Preschool aged siblings (≤ 5 years)*				
≥ 1 siblings	3.8 (49/1,302)	**1.45**	0.96-2.20	**0.0969**
0 sibling	2.6 (39/1,509)	1	-	-
Birth weight (grams)*				
Low birth weight (≤ 2500g)	3.8 (65/1,705)	**1.82**	1.14-2.91	**0.0093**
Normal birth weight (> 2500g),	2.1 (23/1,100)	1	-	-
Birth weight for gestational age				
≤ 10^{th} percentile	4.1 (12/296)	1.33	0.74-2.43	0.4846
>10^{th} percentile	3.0 (76/2,509)	1	-	-
Infant sex				
Female	3.3 (43/1,303)	1.10	0.73-1.67	0.7115
Male	3.0 (45/1,508)	1	-	-
Passive household smoke exposure (Mother's smoking status)*				
Yes	5.3 (34/643)	**2.10**	1.36-3.18	**0.0054**
No	2.5 (51/2,007)	1	-	-
Household crowding (≥ 5 individuals per household)*				
Yes	4.23 (21/497)	**1.56**	0.95-2.54	**0.1488**
No	2.71 (60/2,216)	1	-	-
Socioeconomic status (SES)*				
Highest	4.3 (27/636)	1.26	0.70-2.26	0.5275
Upper-middle	3.7 (20/535)	1.10	0.59-2.07	0.8734
Middle	2.6 (13/499)	0.77	0.38-1.56	0.5955
Lower-middle*	1.6 (9/572)	**0.47**	0.21-1.03	**0.0875**
Lower	3.4 (18/536)	1	-	-
Breast feeding*				
No	3.9 (39/1001)	**1.46**	0.96-2.22	**0.1096**
Yes	1.7 (47/1765)	1	-	-

Variable	*% RSV-H*	*URR*	*95% CI*	*P-value*
Apgar score at 1 minutes				
Low score (0-6)	3.7 (15/397)	1.24	0.72-2.14	0.5619
Normal score (7-10)	3.0 (73/2,401)	1	-	-
Apgar score at 5 minutes				
Low score (0-6)	2.8 (1/36)	0.88	0.12-6.15	0.7082
Normal score (7-10)	3.15 (87/2,760)	1	-	-
Mode of delivery				
Vaginal	3.3 (55/1,649)	1.17	0.77-1.80	0.5212
Cesarean section	2.8 (33/1,162)	1	-	-
Maternal age				
≤25 years old	4.0 (27/681)	1.38	0.89-2.16	0.2267
>25 years old	2.9 (61/2130)	1	-	-
Number of fetuses				
Multiple (twins, triplets)	3.2 (22/681)	1.04	0.65-1.68	0.9639
Single	3.1 (66/2130)			
PPRoM*				
No	3.6 (66/1819)	**1.64**	1.02-2.63	**0.0376**
Yes	2.2 (22/992)	1	-	-
Antenatal steroids				
No	3.0 (55/1816)	0.91	0.59-1.40	0.7626
Yes	3.3 (33/995)	1	-	-
Chronological age (CA) at the start of RSV season*[¥]				
31to ≤52 weeks	4.6 (53/1151)	**2.2**	1.43-3.32	**0.0007**
≤ 30 weeks	2.1 (35/1660)	1	-	-
Student t-test: CA on December 1st (weeks)*[++]				
Outcome	***Mean (95% CI)***	***SD***	***Min-Max***	***P-value***
RSV-H	31.44 (28.0-34.9)	14.7	0-52	0.0027
No RSV-H	26.07 (25.5-26.6)	16.1	0-52	

Note: Risk factor analysis cohort; 88 infants with RSV-associated LRTI admitted to hospital and 2,723 with no RSV-associated LRTI hospitalization during follow up.

-URR - unadjusted relative risk.

*****Indicates significant predictive variable as covariates (p-value ≤ 0.15) in the bivariate analysis.

[¥] To estimate RR for CA was transformed into a categorical (based on the results from the t-test), variable for the univariate analysis, however was maintained as a continuous variable for the logistic regression.

[++]Student's t-test comparing CA (means) on December 1st of the two groups (RSV-H, n=88, vs. no RSV-H, n=2,723).

Table 5-5: Full Model with the Three Independent Predictors Associated With the Risk of Hospitalization after the Stepwise Multivariate Logistic Regression Analysis

Predictive Variable	OR Estimate	(95% CI)	p-Value ≤ 0.05
Household crowding (≥ 5 individuals)	1.68	0.99-2.83	0.0538*
Born in December, January or February	2.58	1.62-4.09	<0.0001
Passive household smoke exposure	2.26	1.42-3.59	0.0006

Note: Logistic regression analysis; 76 infants with RSV-associated LRTI admitted to hospital and 2,407 with no RSV-associated LRTI hospitalization.

*Crowding is just shy of the $p \leq 0.05$ and therefore was included for analysis.

Table 5-6A: Significant Interaction Between Passive Household Smoke Exposure and Birth in December, January or February in Association with RSV-H

Passive Household Smoke Exposure and Born in December, January or February				
Passive household smoke exposure	**No**		**Yes**	
Born Dec.-Jan.	**Yes**	**No**	**Yes**	**No**
RSV-H	**N (%)**	**N (%)**	**N (%)**	**N (%)**
No	429 (93.9)	1527 (98.5)	149 (93.7)	460 (95.0)
Yes*	**28 (6.1)**	**23 (1.5)**	**10 (6.3)**	**24 (5.0)**

* Demonstrates the statistically significant interaction effect ($p<0.0001$) between passive smoke exposure and birth in December, January or February in relation to the RSV-H.

Table 5-6B: Significant Interaction Between Apgar Score at 1 min and Passive Smoke Exposure in Association with RSV-H

Apgar Score at 1 min and Passive Household Smoke Exposure				
Apgar Score at 1 min	**Low score (0-6)**		**Normal score (7-10)**	
Passive household smoke exposure	**Yes**	**No**	**Yes**	**No**
RSV-H	**N (%)**	**N (%)**	**N (%)**	**N (%)**
No	77 (93.9)	280 (96.6)	529 (94.8)	1668 (97.6)
Yes*	**5 (6.1)**	**10 (3.5)**	**29 (5.2)**	**41 (2.4)**

*Demonstrates the statistically significant interaction effect ($p=0.0065$) between Apgar score at 1 min and passive smoke exposure in relation to the RSV-H.

Table 5-6C: Significant Interaction Between Birth Weight for Gestational Age and Passive Smoke Exposure in Association with RSV-H

Birth Weight for Gestational Age and Passive Household Smoke Exposure				
Birth Weight for GA	**Low (≤ 10th percentile)**		**Normal (> 10th percentile)**	
Passive smoke exposure	**Yes**	**No**	**Yes**	**No**
RSV-H	**N (%)**	**N (%)**	**N (%)**	**N (%)**
No	75 (93.8)	187 (96.9)	532 (94.8)	1766 (97.5)
Yes*	**5 (6.3)**	**6 (3.1)**	**29 (5.2)**	**45 (2.5)**

*Demonstrates the statistically significant interaction effect ($p=0.0094$) between birth weight for GA and passive smoke exposure in relation to the RSV-H.

Table 5-7A: Model Fit Statistics for the Three Variable Model

Three-Variable-Model			
Criterion	***Intercept Only***	***Intercept and Covariates***	
AIC	699.956	672.305	
SC	705.802	683.997	
-2 Log L	697.956	668.305	
Test	***Chi-Square***	***DF***	***Pr > ChiSq***
Likelihood Ratio	29.6508	3	<0.0001
Score	33.5894	3	<0.0001
Wald	31.3574	3	<0.0001
R-square: 0.0115		**Max-rescaled R-square:** 0.0483	

Note: Degrees of Freedom (DF).

Table 5-7B: Model Fit Statistics for the Three Variable Model + Interaction

Three Variable Model + Interaction			
Criterion	***Intercept Only***	***Intercept and Covariates***	
AIC	699.956	669.476	
SC	705.802	681.169	
-2 Log L	697.956	665.476	
Test	***Chi-Square***	***DF***	***Pr > ChiSq***
Likelihood Ratio	32.4793	4	<0.0001
Score	34.1601	4	<0.0001
Wald	31.4638	4	<0.0001
R-square: 0.0126		**Max-rescaled R-square:** 0.0528	

Note: Degrees of Freedom (DF).
-Interaction = Passive household smoke exposure and infants born in Dec.-Jan.

Table 5-8: Predictive Variable Scores for RSV-H in the Nova Scotian 32-35 wGA

Predictive Variable	**RSV-H β Coefficient**	**Relative RSV-H β Coefficient**	**Score**
Household crowding (≥ 5 individuals)	0.255	0.224	22
Born in December, January or February	0.474	0.417	42
Passive household smoke exposure	0.408	0.359	36
TOTAL:	-	1	100

Note: Logistic regression analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization.

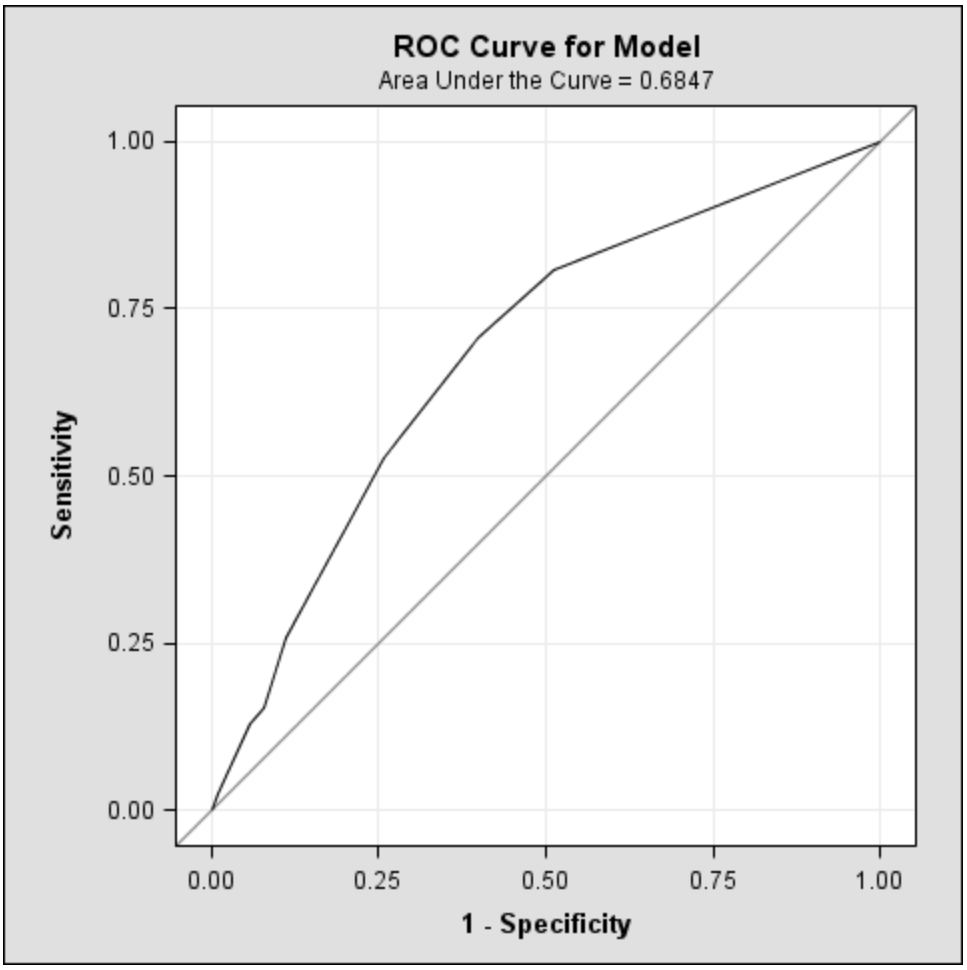


Figure 5F-1: Receiver operating characteristic analysis for RSV-H scoring tool for the three variable predictive model

Note: Sensitivity (the proportion of true positives results) is shown on the y axis, going from 0-1 (0-100%) and 1-specificity (the proportion of false positives) is sowing on the x-axis, going from 0-1 (0-100%).

Table 5-9: Receiver Operating Characteristics, Probability Level, Sensitivity, 1-Specificity and Associated RSV-H Score (0-100)

RSV-H Probability level	TP	TN	FP	FN	Sensitivity	1-Specificity	RSV-H Score (0-100)
0.135	2	2,455	23	76	0.036	0.009	100
0.086	10	2,338	140	68	0.128	0.056	78
0.065	**12**	**2,282**	**196**	**66**	**0.154**	**0.079**	**64**
0.057	20	2,200	278	58	0.256	0.112	58
0.040	41	1,842	636	37	0.526	0.257	42
0.035	**55**	**1,494**	**984**	**23**	**0.705**	**0.397**	**36**
0.026	63	1,211	1,267	15	0.807	0.511	22
0.016	78	0	2,478	0	1	1	0

Note: Event=RSV-associated hospitalization (RSV-H), Nonevent=No RSV-associated hospitalization (no-RSV-H).
-Predictive tool analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization. There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool, therefore were excluded for the final RSV-H analysis.
-Sensitivity (rate of hospitalized infants correctly predicted, or "true positives ratio") and specificity (rate of non-hospitalized infants correctly predicted or "true negative ratio").

Table 5-10: Number and Percentage of RSV-H by Risk Group and Score Range

Risk group	Score range	RSV-H n (%)	No RSV-H n (%)	Total
Low risk	0-41	23 **(1.5%)**	1,494 (98.5%)	1,517
Moderate risk	42-64	43 **(5.2%)**	788 (94.8%)	831
Highest risk	65-100	12 **(5.8%)**	196 (94.2%)	208

Note: Predictive tool analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization. There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool, therefore were excluded for the final RSV-H analysis.

Table 5-11: Group Categorization, the Mean Predicted Probabilities and Relative Risk Estimates for RSV-Associated Hospitalization

Risk group Categorization	Score range	Mean predicted probability of RSV-H	Relative risk estimates (comparing means between low risk level)
Low risk	0-41	0.0212	-
Moderate risk	42-64	0.0433	2.04
High risk	65-100	0.0860	4.05

Note: Predictive tool analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization. There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool, therefore were excluded for the final RSV-H analysis.
-Low risk n = 1,879; moderate risk group n = 469; and high risk group n = 208.

Table 5-12: Test Properties of RSV-Associated Hospitalization

Scoring Tool Results:	**RSV-H**		**No RSV-H**		**LR+**	**+ Post Test Probability**
	N	TP/(TP+FN)= **Sensitivity**	**N**	TN/(FP+TN)= **Specificity**		
High Risk	12	12/78= 0.154	196	196/2,478= 0.079	**1.95**	**6.2%**
Moderate Risk	43	43/78= 0.551	440	788/2,478= 0.318	**1.73**	**5.5%**
Low Risk	23	23/78= 0.295	1,842	1,494 /2,478= 0.603	**0.49**	**1.6%**
Total:	**78**	**-**	**2,478**	-		

Note: Predictive tool analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization. There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool, therefore were excluded for the final RSV-H analysis.
-LR=Likelihood ratio; LR+ = sensitivity/ (1-specificty).
-Pre-test rate of RSV associated hospitalization = 3.1% (88/2,811).
-Positive post-test probability; calculated from post-test odds ((pre-test odds)*(LR+)).

Table 5-13: Results of the RSV-H Scoring Tool for the 41 Point Cut-Off Between Moderate and High Risk (>42) and Low Risk (0-41):

Scoring Tool Cut Off: ***Moderate and High risk (>42) and Low risk (0-41)***	**RSV-H (+)**	**No RSV-H (-)**	**Total**
Scoring Tool + (Positive predictive outcome; RSV-H)	TP = 55	FP = 984	1,039
Scoring Tool – (Negative predicted outcome; no RSV-H)	FN = 23	TN = 1,494	1,517
Total	78	2,478	2,556
Sensitivity **Specificity** **PPV** **NPV**	71% 60% 5% 98%	**Likelihood Ratio +** **Likelihood Ratio –** **Positive Post Test Probability**	**1.78** 0.56 **5.5%**

Note: Predictive tool analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization. There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool, therefore were excluded for the final RSV-H analysis.
-Pre-test probability for infants = 3.1% (78/2,556).
-Pre-test odds = 0.031/(1-0.031)=0.032.
-TP (# of correctly predicted events); TN (# of correctly predicted non-events); FP (# of non-events predicted as events); FN (# of events predicted as non-events).
-Sensitivity = TP/(TP + FN).
-Specificity = TN/(TN + FP).
-PPV= TP/(TP+FP).
-NPV= TN/(TN+FN).
-LR=Likelihood ratio; LR+ = sensitivity/ (1-specificty), LR- = (1-specificty)/sensitivity.
-Pre-test rate of RSV associated hospitalization = 3.1% (88/2,811).
-Positive post-test probability; calculated from post-test odds ((pre-test odds)*(LR+)).

Table 5-14: Results of the RSV-H Scoring Tool for the 64 Point Cut-Off Between High Risk (65-100) and Low and Moderate Risk (0-64):

Scoring Tool Cut Off: ***High risk (65-100)*** ***and*** ***Low and Moderate risk (0-64)***	**RSV-H (+)**	**No RSV-H (-)**	**Total**
Scoring Tool + (Positive predictive outcome; RSV-H)	TP = 12	FP = 196	208
Scoring Tool – (Negative predicted outcome; no RSV-H)	FN = 66	TN = 2,282	2,348
Total	78	2,478	2,556
Sensitivity **Specificity** **PPV** **NPV**	15% 92% 5.8% 97%	**Likelihood Ratio +** **Likelihood Ratio –** **Positive Post Test Probability**	**1.88** 0.95 **5.8%**

Note: Predictive tool analysis; 78 infants with RSV-associated LRTI admitted to hospital and 2,478 with no RSV-associated LRTI hospitalization. There were 255 infants that had missing values for the three predictive variables included in the RSV-H scoring tool, therefore were excluded for the final RSV-H analysis.
-Pre-test probability for infants = 3.1% (78/2,556).
-Pre-test odds = 0.031/(1-0.031)=0.032.
-TP (# of correctly predicted events); TN (# of correctly predicted non-events); FP (# of non-events predicted as events); FN (# of events predicted as non-events).
-Sensitivity = TP/(TP + FN).
-Specificity = TN/(TN + FP).
-PPV= TP/(TP+FP).
-NPV= TN/(TN+FN).
-LR=Likelihood ratio; LR+ = sensitivity/ (1-specificty), LR- = (1-specificty)/sensitivity.
-Pre-test rate of RSV associated hospitalization = 3.1% (88/2,811).
-Positive post-test probability; calculated from post-test odds ((pre-test odds)*(LR+)).

CHAPTER 6: DISCUSSION

6.1 MAJOR FINDINGS

6.1.1 Risk Factors Attributable to RSV-Associated Hospitalization

This thesis aimed to identify risk factors in Nova Scotian infants 32-35 completed wGA that were associated with RSV-H during the infant's first year of life. The multivariate logistic procedure included 76 infants with RSV-H and 2,407 infants with no RSV-H due to missing values for the risk factors included in the final scoring tool. Independent predictors for RSV-H were infants born in December, January or February, passive household smoke exposure and household crowding.

Birth Month

Birth month was defined as infant's birth in December, January or February and was compared to infants born March-November. Logistic regression analysis indicated the highest odds of hospitalization (OR=2.58; 95% CI=1.62-4.09; $p<0.0001$) associated to Dec.-Jan. birth month. In the PICNIC study, birth between November and January was one of the five independent predictors for the increased risk of RSV hospitalization [11] (Table 2-3). Additionally, a study conducted in Germany concluded that birth discharge between October and December was associated with a higher risk of hospitalization (OR: 2.1; 95%CI: 0.99-4.4) [56].

In the Northern hemisphere, RSV epidemics peak in February and March. In Canada, RSV season usually begins between November and January and persists for four to five months as stated by the Public Health Agency of Canada [97]. Appendix C shows the number of RSV positive laboratory isolates per month in Canada between 1997 and 1998 as reported by the Notifiable Disease Reporting System, Division of Disease Surveillance, Centre for Infectious Disease Prevention and Control, Health Canada. The figure indicates that the highest associating rates of RSV begin in December, with the highest peak occurring in January and February and eventually tapering off in April [98]. RSV season varies across Canada, but by using NS data, we were able to view the

province's RSV seasonal variation in detail. Our data indicated that the peak of RSV hospitalization in NS between 1998 and 2008, among infants 32-35 wGA, occurred in January and February (~6.5% RSV-H; Table 5-3).

In an extensive review of the literature, Simoes et al. conclude that birth during the first half of the RSV season is a risk factor for both the development of RSV LRTI as well as RSV hospitalization [75]. Birth during the peak months of an RSV seasonal outbreak would increase the frequency of exposure to the virus at a very young age. Simoes *et al.* also state that maternal antibodies titers to RSV show a seasonal variation with lower levels within the first half of the RSV season due to the fact that mothers have not been exposed to RSV.

The two Spanish Flip and Flip-2 studies, however, demonstrated that infants with chronological age < 10 weeks at onset of RSV season were at higher risk for hospitalization. In our logistics regression analysis, CA was not found statistically significant with RSV-H. This is indicative that infants born specifically during the peak months of RSV activity in NS (Dec.-Jan. between 1998-2008) are most susceptible to severe infection.

Passive Household Smoke Exposure

Mother's smoking status at the time of delivery was used as a proxy measurement for passive household smoke exposure. Infants with passive smoke exposure demonstrated an OR associated with RSV-H of 2.26 (95% CI=1.62-4.09; $p<0.0006$).

A recent systematic review on the role of environmental tobacco smoke (ETS) exposure as a risk factor for serious RSV disease among infants and young children concluded that there was ample evidence that ETS increases the risk of RSV-attributable hospitalization [69]. The review highlights a number of different study designs (retrospective, prospective, case-control) among different populations (premature infants, general population). In most studies, data on the child's ETS exposure level were gathered from parent or caregiver report, whether through direct data collection for the

study or, through retrospective review of the mother's prenatal health records. The studies compare different definitions for ETS exposure (maternal smoking during pregnancy, parental smoking at home, ≥ 2 smokers in the household) as well as different definitions for the outcome measurements (RSV LRTI, RSV hospitalization, length of hospital stay).

Among the studies in premature infants, 5 studies assessing laboratory-confirmed RSV [11, 20, 21, 65, 68] found ETS exposure to be a significant risk factor in either bivariate or multivariate analysis. Significant results in the multivariate analysis were seen in the prospective PICNIC study which defined smoke exposure as ≥ 2 smokers in the household (OR 1.71; 95% CI 0.97-3.00; P= 0.064) while the FLIP-2 study classified smoke exposure as maternal smoking during pregnancy (OR=1.61; 95% CI=1.16 –2.25; p=0.0044). In comparing our study to the PICNIC and FLIP-2 study, there were variations in data collection, definition of ETS or the potential misclassification of RSV disease (ICD-CM codes vs. laboratory testing for RSV). These differences may produce a bias leading to either an over or underestimate of the true ETS exposure risk or to a null finding. In particular, the difference in definitions and data collection (prospective vs. retrospective) for smoking exposure between our study and that of the prospective PICNIC study would elucidate the variation in ORs within the Canadian population.

Regardless of the variation in strength, the evidence presented for the association between ETS and severe RSV infection is convincing because different patient populations in a variety of countries and cultures have been studied using different methodology [69]. Our findings provide additional rationale for the health promotion and prevention of smoke exposure, especially high-risk groups, such as premature infants who are at an increased risk for serious RSV disease.

Household Crowding

Household crowding was defined as ≥ 5 individuals per household counting the study subject. The number of previous deliveries in proceeding 18 years (gravida) and

marital status (single vs. married) was used as a proxy for the variable. In the multivariate analysis, household crowding demonstrated an OR=1.68 (95% CI=0.99-2.83; p=0.05).

Residential crowding can facilitate the spread of RSV infection through viral shedding under close contact conditions, increasing the likelihood of exposure to the virus. A recent systematic review assessed the association between crowding in the home and the increased risk of severe RSV disease in children (< 5 years old). Among all studies, crowding was measured in numerous ways; number of residents per household, room or square meter. Others measured the number or age of children or siblings in the home or how many individuals shared the subject's bed. Most of the studies examined laboratory-confirmed RSV hospitalization outcome (other outcome measurements included confirmed diagnosis and treatment in the in the emergency department, outpatient clinic, hospital or pediatricians office) [81] . Significant effects > 1 were consistent across different populations (i.e. native race, premature infant, and the general population), study design (case control, prospective, retrospective) and across multiple countries [81].

Exposure to RSV disease in a crowded environment is of concern for premature infants who have an augmented risk of developing more severe disease. The case-control study by Figueras-Aloy et al. found an association with crowding (defined as four or more household inhabitants, not including school-aged siblings or the study subject). However, this variable was not significant in the larger FLIP-2 study. The PICNIC study found that infants in homes with five or more residents were significantly at risk of hospitalization due to RSV. These results were comparable in the magnitude of effect size to our study (OR=1.69 vs. 1.68 respectively), despite the difference in methodological design, data collection and definition of crowding and outcome measurement (ICD-CM codes vs. laboratory testing for RSV).

6.1.2 Significant Interaction Effects Between Predictive Variables in Association with RSV-Associated Hospitalization

Interpreting the independent contribution of each risk factor attributable to RSV-H becomes challenging when certain variables closely relate to one another in the multivariate analysis. Effect modification occurs when the effect measure (i.e. RR, OR) depends on the level of another factor (also known as interaction effect). In other words, two risk factors are considered to be independent if their effects are additive, and they are considered to *interact* if their joint effect is different from the sum of their independent effects. When interpreting the interaction of two risk factors on RSV-H, we used a multiplicative model (i.e RR measure) to adjust for all other potential confounders. We calculated and interpreted both the independent and joint effects of the risk factors under consideration using relative risk measurement.

Significant Interaction Effects

Birth in December-February as well as exposure to tobacco smoke were the most significantly independent predictors of RSV hospitalization, as well, demonstrated the most significant interaction effect ($p<0.0001$) (Table 5-6b) in the univariate analysis. Results indicate the relationship between birth in Dec.-Jan. and exposure to smoking jointly increasing the risk of RSV hospitalization. To our knowledge no study has interpreted the combination effect on birth month and smoking. The same can be said for the interaction and combination effect between passive household smoke exposure and both low Apgar score at 1 minute ($p=0.0065$) as well as weight for gestational age ($p=0.0094$). Independently, small for gestational age has only otherwise been examined and found significant in the Canadian PICNIC study. This is the first study to look at the relationship between Apgar score and RSV-H. The potentially modifiable smoking variable in addition to the non-modifiable risk factors (birth month, low apgar score a low weight for GA) indicated heightened risk of RSV-H in the univariate analysis.

Interaction Effects in the Three Variable Model

The inclusion of these significant interactions within the three variable predictive model was an attempt to control for any multipliable interaction effects between risk factors. In addition, to further interpret the question, “Do the non-modifiable risk factors

depend on smoking in relation to RSV-H outcome?" Smoking and birth month was the only interaction that remained significant in the multivariate analysis.

In any observational study, it is difficult to assess the direct source of ETS exposure. The variation in definition of ETS as well as the different methods of collecting the data may introduce unknown confounders as well as reporting bias. In addition, ETS exposure can come through many avenues, making it hard to find the direct source of measurement. Examples include in utero exposure through active and passive maternal smoking and postnatal exposure from parents, caregiver and other individuals living at the home. To further complicate things, ETS exposure can also occur in public places such as day care settings and other residences of friends and family. Lastly, ETS smoke exposure is known to interact closely with other variables, however the causal mechanism underlying the epidemiologic association between ETS and the severity of RSV infection is unclear [69, 99]. One can assume however, that regardless of the direct source, an infant's level of exposure to ETS would be elevated when situated in a closed indoor space in comparison to the open outdoor environment. Environmental air pollutants such as tobacco smoke and ambient air pollutants make children more susceptible to acquiring an RSV infection. The more pollutants a child is exposed to increases the likelihood of a severe infection [98, 99]. In terms of the interaction effect, we can hypothesize on the increased frequency of ETS exposure for infants born in the colder winter months in NS, who would spend a greater deal of their early life (when most susceptible to infection) indoors in comparison to infants born in warmer months.

6.1.3 Respiratory Syncytial Virus Scoring Tool

The purpose of the RSV-H scoring tool was to develop a clinical instrument based on risk factors present at birth that would discriminate between infants at high, moderate and low risk for RSV-H. Given the high cost of RSV prophylaxis, identifying the highest-risk infants would help to target local health interventions and cost effective strategies. The model assigned individual weights (based on β coefficients) to the three significant risk factors. Each risk factor was identified as being either present or absent, with a positive answer being assigned a pre-determined score between 22 and 42, while a

negative answer was assigned a score of zero. The final score was the sum of the individual scores ranging from 0-100 (Table 5-8). ROC analysis was conducted in order to discriminate between infants with low, moderate and high risk of RSV-H.

We tabulated 1) the predicted probability of an RSV-H event, 2) the sensitivity of the model, 3) the specificity of the model and 4) the AUC of the overall model. Highest sensitivity and lowest false positive rates identified by the RSV risk scoring tool provided a cut-off value of 41, between low risk infants (0-41) and moderate and high risk infants (>42). The tool determined a second cut-off score of 64 between moderate (42-64) and high risk infants (65-100). In other words, the scoring tool distinguishes between low risk (infants with a 1.5% RSV-H), moderate risk (5.2% RSV-H) and high risk infants (5.8% RSV-H) (Table 5-10).

Comparing the results from the 2x2 tables (Table 5-13, 5-14) indicate that the determination of the cut off between low and moderate risk and high risk infants (64 point cut-off), decreased the scoring tools sensitivity (71% vs. 15% respectively), however, it increased its specificity (60% vs. 92% respectively). Results indicate that the multifactorial scoring tool for RSV-H may not be valid for determining the proportion of truly hospitalized infants (TP) in the highest risk category, those infants 32-35 wGA who would most benefit from Pz prophylaxis.

Predictive Values of the RSV-H Scoring Tool

The PPV calculations were very similar when comparing the results from the 2x2 tables; 5 vs. 5.8% for the 41 and 64 point cut-off, respectively. Infants classified in the higher risk group did not necessarily have an overall higher probability of RSV-H in comparison to those with lower scores. The LR was calculated in order to determine whether knowing the results of the scoring tool (post-test probability) would improve on the pre-test probability of RSV-H. In general, the higher a LR+, the greater the difference between pre and post-test probabilities of the stronger it is in predicting outcome. Conversely, the smaller the LR–, the better it is in predicting the absence of the outcome. LR+ of 1 to 2 and negative LR– of 0.5 to 1 alter probability to a small (and rarely

important) degree in clinical practise [100]. Our LRs+ were very similar between both cut-off points (LR+1.78 vs. LR+ 1.88, Tables 5-13 and 5-14). The LRs indicate that the post-test probability is slightly greater than the pre-test probability at both cut-off points, however, overall the scoring tool adds little additional value for clinicians who wish to identify infants at highest risk (65 point cut off).

The sensitivity and specificity of the PICNIC predictive tool were 68% and 72% respectively. The FLIP study found a 62% sensitivity and 66% specificity respectively in the study's validation of the predicting tool [11, 21]. Several other models have been developed and validated [63, 64, 101, 102]. These models, also based on logistic regressions, varied between 8-6 variable models [63]. Sensitivity of the models ranged from 0.51-0.88 and specificity ranged from 0.63-0.72 [49]. PPV % ranged from 13-75%, NPV % ranged from 73-99% while LR for a positive outcome (RSV-H) ranged from 1.46-3.17 among the European models. It is important to note that these models optimized which variables would discriminate between only two groups; RSV hospitalized vs. non RSV-hospitalized premature infants (33-35 wGA). Our results lie in between these ranges when distinguishing between low and moderate/higher risk infants, still, when attempting a second cut-off, our tool was unsuccessful at predicting between moderate and the highest risk infants.

Post-Test Probabilities of the RSV-H Scoring Tool

By multiplying the pre-test odds by the LR (converting between odds and probabilities, section 4.4.4), the post-test probabilities confirmed similar results and provides further insight on the tools validity in clinical practice. The post-test probabilities were very similar between both 41 and 64 points cut-off (5.5% vs. 5.8%, respectively). Overall, this indicates the scoring tools inability to target those at the highest risk for RSV-H. Additionally, the post-test probabilities are not very different than the pre-test probability of 3.1%, decreasing its overall validity for clinical application.

6.2 HEALTH CARE PRACTICE IMPLICATIONS OF THE SCORING TOOL

To be helpful in a clinical practice, a predictive model must discriminate between infants who will be hospitalized and those who will not be hospitalized. Although our scoring tool accurately classified lowest risk infants apart from moderate/high risk infants, it failed to accurately predict infants at the highest risk for RSV-H. This is particularly important given the high cost of palivizumab prophylaxis for treatment and prevention of severe RSV disease.

A multifactorial scoring tool for RSV-H estimated from information collected from administrative databases may not be valid for targeting premature infants 32-35 completed wGA at highest risk RSV-H. However, our findings will add to the current literature for infants in the 32-35 wGA, specifically in NS. Risk factors associated with more severe illness associated with RSV were identified among 32-35 wGA among premature infants in NS; passive household smoke exposure, household crowding and birth in December, January and February. Understanding an infant's risk based on both modifiable and non-modifiable risk factors can result in appropriate health preventative strategies as well as provide evidence based information to parents/caregivers. All individuals should be provided this information and educated on what they can do to help prevent severe RSV infection.

Our study indicated that infants that are exposed to household smoke are at a much higher risk for RSV-H. The prevalence of maternal smoking in Canadian mothers, at 13.4%, has been gradually declining. On the other hand, the rate in NS mothers, at 23.6%, has increased over the past decade [103]. In addition to the alarmingly high rates of maternal smoking in the province, the results from this study should encourage health professionals to prioritize educational practices on this modifiable risk factor. Mothers who continue to smoke in the home should be made aware of the risk associated between smoking and RSV-H, especially if their child is at high risk (i.e. premature).

Although birth month is a non-modifiable risk factor, the cumulative effect of multiple risk factors in this study has shown an increased risk for RSV-H and should be

considered during the decision making process for receipt of Pz. Premature infants may be more vulnerable to the negative impact of household crowding, such as increased exposure to RSV through contact with more people. Caregivers should be educated on frequent hand washing and wiping of hard surfaces with soap and water or disinfectant which may help stop infection and spread of RSV, particularly in high density households.

For the NS population, the development of the RSV-H scoring tool provided further insight on modifiable risk factors such as smoking as well as non-modifiable risk factors such as household crowding and birth during the peak months of the province's RSV season (in addition to 32-35 weeks GA). Additionally, the combination of ETS with certain non-modifiable risk factors (i.e. birth during Dec-Jan.) may pose an even greater risk of severe RSV infection. In terms of clinical use, the RSV-H scoring tool could be used as an educational guideline for health professionals, outlining the importance of both modifiable and non-modifiable risk factors to parents and caregivers, specifically for infants born 32-35 weeks GA.

Of importance is the awareness that the assessment of risk factors for this study was applied primarily to a cohort of infants defined by a set of clinically relevant variables, rather than directly to individuals. In a clinical practice, the overall risk-benefit assessment for Pz administration would be case by case depending on the clinical setting. Therefore, our risk assessment for the NS population serves only as a guideline and should not be taken as an absolute definition of high risk. Several promising risk factors (e.g. child care attendance) remain to be investigated for this patient population. The infants or siblings of the infant attending day cares bring an at-risk child into close proximity with potentially infected individuals. Day care attendance had the largest independent impact on hospitalizations for RSV in the PICNIC study. Additionally, the AAP suggests child care attendance as one of the two risk factors used in their recommendation for RSV prophylaxis. Our study was unable to include child care attendance as a predictive variable for analysis. While the choice of predictors for this

study was restricted to variables available in the administrative database, future cohort studies should consider including these important predictors in their study design.

There are certain inconsistencies throughout the studies attempting to define risk factors associated with RSV-H, however, other Canadian studies have underlined the importance of day care attendance as a risk factor for RSV-H. It is important that local data in addition to the current recommendations by both the CPS and AAP be used as a guideline when providing care to patients at risk for RSV-H. For these infants, major risk factors including birth month, ETS exposure, living conditions, including crowding, and day care attendance should be included when making educated decisions regarding optimal care.

6.3 STRENGTHS AND LIMITATIONS

To the author's best knowledge, this was the first retrospective cohort study conducted in Canada that examined information present at baseline for premature infants born 32-35 wGA for ten consecutive RSV season years. The 10 RSV seasons provided inclusion of 88 infants hospitalized for RSV compared to the smaller 66 hospitalized cohort studied in the previous Canadian PICNIC study for analysis [11]. In addition, the extensive information present in the Atlee database, allowed this study to examine variables associated with RSV-H that have not been previously studied in Canada; Apgar score, SES, PPRoM and antenatal steroids.

Despite the strengths, our study has a number of limitations. Due to its retrospective nature, analyses of certain predictive variables potentially associated with RSV hospitalization were not available. For example, we could not assess information on daycare attendance, family history of allergies/ respiratory disease, low maternal RSV antibody levels, and race/ethnicity, all of which the literature suggests are likely contributed to RSV hospitalization. We were only able to provide proxy variables for the presence of preschool aged siblings, household crowding, and SES. The assumption is made when looking at crowding or household environment, assuming that marital status is unchanged and all infants and previous births aged <18 at the time the baby was born still live in the household. Looking at socioeconomic status, family income was estimated using neighborhood income from census data, and factors related to social status were unavailable. While census data have been used as a proxy for income in other studies, this measure cannot be interpreted as if it were collected at the individual level [104].

The percent of missing values in our study was relatively low, therefore the effect of the missing data was not believed to have an overall direct impact or lead to any bias in our results. Certain variables were self-reported by the mother, such as smoking status, which introduces the important risk of reporting bias among the predictive variables.

We also determined three statistically significant interaction effects that appeared to slightly increase our three variable model's predictive capability through stepwise logistic analysis. A statistical interaction occurs when the effect of one independent variable on the dependent variable changes depending on the level of another independent variable. The two variables interacting, passive household smoke exposure and birth in Dec.-Jan., lead to results that were not anticipated on the basis of the main effects (the effect of each single risk factor associated with RSV-H, ignoring the effect of all other independent variables). Our inability to precisely control for these potentially multi-factorial associations in our final model may have contributed to bias leading to either an over or underestimate of the strength of association between the predictive variables under investigation and RSV-H.

The use of the CIHI's ICD-CM codes as our outcome measurement for estimates of RSV-H is subject to misclassification. True cases of RSV may not have been captured because of failure to conduct a laboratory test of the pathogen. This prohibited us to accurately determine whether laboratory confirmation of RSV was available, and if available whether or not it indicated a positive RSV result. The study makes an assumption that those health care professionals who make the final diagnosis for RSV specific disease had reason to believe that RSV was the cause for illness. Furthermore, reliance on RSV-specific ICD-CM diagnoses is subject to an unknown magnitude of error, including incomplete information and miscoding. If by example true RSV infected patients were coded as unspecified pneumonia or bronchiolitis instead of RSV pneumonia (J12.1) or acute bronchiolitis due to RSV (J21.0), the RSV-coded diagnoses would be more likely to be underestimated in the patient discharge[105]. Lastly, a study compared virology reports from laboratory-confirmed RSV cases with the primary and secondary diagnostic code in the patient's records between October, 1996, and May, 1999. The RSV-specific diagnostic code was used for only 77% of the laboratory-confirmed cases [105]. These findings suggest that the actual number of cases could be substantially greater. If it the number of RSV-H cases reported in Table 2 were indeed underestimated, the actual cases of RVS-H would be greater than the estimates reported here. If indeed there were higher annual number of RSV-H in the province, additional

analysis on certain risk factor associated to RSV-H could have provided a more accurate depiction of the overall effect of risk factors. In order to illustrate the change in overall effect size of risk factors, the OR (the odds of the outcome in one group divided by the odds of the outcome in the other group) is interpreted. Here, p_1 refers to the probability of the outcome in group 1, and p_2 is the probability of the outcome in group 2.

$$\text{Odds Ratio} = \frac{\rho_1/(1-\rho_1)}{\rho_2/(1-\rho_2)}$$

One could hypothesize that an increase in the probability of RSV-H in group 1, would increase the overall effect size (OR) for that particular risk factor.

A study that tested the accuracy of ICD-CM codes in detecting community-acquired pneumococcal pneumonia for incidence and vaccine efficacy concluded that without the ability to validate diagnoses by review of medical charts, researchers must carefully decide which codes would minimize misclassification. The authors go on to suggest that the selection of the most specific ICD code or group code of ICD-CM codes to identify the disease under investigation should be conducted. Table 4-1 provides the description on the ICD-CM codes selected for this study. Only the *RSV-specific* diagnosis codes were chosen for analysis. Narrowing the definition for RSV by selecting only RSV-specific codes was elected as the most accurate definition for the required outcome definition.

The application of a retrospective study design invites discussion of methodological limitations. The major weakness of retrospective data is that they are often incomplete. For example, information on infant data using the NSPAD as well as the CIHI's data on outcome measurement is not a true depiction of the 32-35 wGA cohort in NS. Infants whose clinical diagnosis of RSV was not confirmed through ICD-CM code would not have been captured in the study cohort. The true rates of potential risk factors and RSV-H could be either under or overestimated (both the numerator and or

denominator that make up the potential effect). Therefore, overall rates may not be a true representation of the actual effect.

The unavailable data of potentially important risk factors and of course our inability to precisely measure our proxy variables, demonstrates the inherently difficult accuracy is assessing data collected in this retrospective nature. Prospective studies such as the PICNIC study, on the other hand, often allow follow up that provide investigators the ability to measure important variables completely and accurately. Our study was however able to include 10 separate RSV seasons for analysis. This longer time period allowed for cases of RSV-H to appear periodically throughout the study, which minimized bias due to seasonal variations, for example. Additionally, in order to collect enough provincial data for significant results, a single RSV season would not suffice in NS.

Observational studies are prone to incomplete data and other unmeasured confounders. There are always certain limitations imposed by how and what data are captured within a given dataset. Among the many large epidemiological studies of RSV hospitalization conducted in Spain, the smaller case-control study by Figueras-Aloy et al. found an association with crowding, however this variable was not significant in the larger 2-cohort study (Flip-2) [20]. Since RSV infections were likely to have been the most severe cases, this might have led to selection bias in constructing the dataset. Additionally, the FLIP study matched variables between the hospitalized cases and the controls. There is always a possibility that matching for known factors will also indirectly match and confound unknown related factors. This may lead to a reduction in the significance of each separate risk factor, leading to spurious results.

Confounding bias is another potential obstacle when determining the association between certain risk factors and RSV-H in observational studies. Interpreting accurate effect measurements from the regression analysis can prove challenging as results may be confounded by inclusions of variables that closely relate to each other. Differentiating independent sources of exposure cannot be defined with certainty. Smoking status as a

well-known association with SES, as well, SES has been linked as a predictive effect of ETS exposure in children [106]. In addition to SES, smoking status is also predicted by race, educational attainment, and marital status [107]. As previously illustrated, tobacco smoke exposure is also a known risk factor for serious RSV disease [69, 99]. Residential crowding likely increases the child's exposure to RSV, but residential crowding also have been identified as a proxy for SES [108]. Smoking during pregnancy increases the risk of adverse outcomes, such as fetal growth restriction and preterm birth [109, 110], which are both risks for severe RSV disease. These are all examples illustrating potential confounders, leading to spurious findings in the overall effect or risk factors in relation to RSV severe disease. This study conducted multivariate analysis to control for potential confounders; as well included statistical modeling incorporating significant interaction effects for further analysis and interpretation. In reporting future studies, greater details about effect modification(s) may aid in assessment of the potential true independent interaction(s) between risk factor(s) and serious RSV disease.

Lastly, the results of this investigation will be based solely on information collected from residents of NS, and although useful for stakeholders throughout this Canadian province, caution should be taken when applying this information to clinical decision-making and public health agencies in other geographic locations. For example, Canadian Inuits tend to live in small houses with a median of six persons per house, with limited natural air exchange, and 94% of the households include tobacco smokers [111]. Dramatically higher rates of hospitalization with bronchiolitis have been described in isolated Inuit populations in Canada: (590 per 1000 infants in the first year of life in the Kitikmeot region of Nunavut between 2000 and 2004). It is likely that approximately one-half of such cases are due to RSV infection [112]. This is not a reasonable comparison to the population of moderately premature infants 32-35 wGA used to provide results for this study.

6.4 FUTURE DIRECTIONS

Further research is required on certain variables within the NS population born 32-35 wGA. Although there is a growing body of evidence suggesting that ETS exposure places infants and young children at increased risk of RSV-H, misclassification of ETS exposure is a continual challenge in studying associations of ETS exposure with disease. There is still a debate regarding the level of exposure and the avenue of ETS exposure, not just whether an infant has been exposed to tobacco smoke [69, 113, 114].

The Atlee database lacked sufficient information regarding daycare attendance, which in addition to preschool age siblings, is used by the AAP in their current recommendation of risk factors for RSV prophylaxis. It may be that certain risk factors are in association with other variables (interaction effect); hence its overall impact in the model is diluted. Specifically, this can be said for low Apgar score and weight for GA in relationship with passive household smoking. Whether certain variables act independently or additively and whether or not these variables warrant inclusion in the model is subject to further investigation. To examine the accuracy of risk factors available at baseline to predict hospitalization due to severe RSV infection among moderately premature infants, further prospective validation of such a design is necessary to determine whether a multifactor prediction tool does indeed target the highest risk infants of this GA in NS.

In addition to further investigation on potentially attributable risk factors for this patient population, it is not clear what recommendation should apply to Inuit infants, to all Aboriginal infants or to all infants in remote communities. A better understanding of the seasonality of RSV in infants in Northern communities remains current CPS research priority. Lastly, as stated by the CPS in the position statement for prevention of RSV infection, further assessment and validation of clinical assessment tools used to identify infants born 32-35 weeks gestation is a required research priority [115].

By using data from provincial databases, this study provides further understanding of the risk factors and their interrelationships for infants born between 32-35 wGA in NS. This information has the potential to improve standards of care by providing evidence based education to parents, caregivers, and communities regarding the importance of reducing both exposure and transmission of RSV, which are critical to RSV prevention. Public health interventions with this educational focus (i.e. significant risk factors) should be targeted toward high-risk groups during periods of peak transmission to prevent the development of disease severe enough to require hospitalization. Understanding an infant's risk can improve health outcomes, the effectiveness of prophylaxis and minimize costs associated with disease. Inexpensive, simple preventative practises, such as good hand hygiene in the home and limiting direct contact of high-risk children with other children and adults with respiratory tract infections, should remain high priority for RSV prevention.

CHAPTER 7: CONCLUSIONS

RSV is the leading cause of respiratory tract infection in infants and young children. Although most infections cause mild disease, RSV is a major cause of hospitalization, especially among premature infants who are at higher risk for serious infection.

We used existing data to investigate whether a readily available set of risk factors in premature infants born 32-35 wGA could be included in a multifactor prediction model for RSV-H. Of these, three of the variables investigated – infants born in December, January or February, passive household smoke exposure and household crowding –were significantly associated with hospitalization. Passive household smoke exposure and infants born in Dec.-Jan. demonstrated the most significant interaction effect associated with RSV-H. In summary, with the exception of determining between low and moderate to high risk infants, our scoring tool failed to accurately predict hospitalization between moderate risk and infants at the highest risk for RSV-H. This was determined by the overall interpretation of our predictive values (PPV, NPV, RL+, LR-, post-test probability) based on calculations from the validity of the scoring tool (ROC; sensitivity, specificity, AUC) across the full range of cut offs.

Passive household smoke exposure demonstrated an association with RSV-H; however, researchers must continue to meticulously investigate ETS exposure with disease, in order to prevent ETS exposure misclassification. As such, further research is warranted in investigating any interaction effect(s) related to smoking, as well any other variable(s) that are believed to be closely related when conducting multivariate regression analysis of risk factors associated with RSV-H. Additionally, our scoring tool was unable to include data on several important measures identified in the literature (i.e. child care attendance), therefore the study should not be considered conclusive. A prospective study of the three identified risk factors in addition to a tested combination of other notable important risk factors is warranted for further study among moderately premature infants in NS. In the development and validation of such a multifactorial scoring tool,

local characteristics should be considered when defining the components making up the model, as well as the definitions of each variables of inclusion. This will help determine the most cost-effective strategy in providing the highest risk infants in the 32-35 wGA palivizumab prophylaxis.

Given that prophylaxis of RSV with palivizumab is not a complete intervention without other preventative measures it is important to continue emphasizing other means of prevention. These include, however are not limited to, eliminating exposure to tobacco smoke as well as continuing education regarding hand washing when siblings or care givers have been exposed to RSV. In clinical practice, our tool could be used as an educational guideline for health care professionals who wish to inform parents and caregivers on both significantly associated RSV-H modifiable and non-modifiable risk factors. For example, if you smoke, do not smoke in the house or car. High frequencies of secondhand smoke may increase the infant's susceptibility to serious RSV associated infections. During times of RSV seasonal outbreaks, try to avoid the infant's exposure to people with respiratory infections, such as colds. Pay attention to wiping off surfaces with disinfectants, and avoid sharing cups or toys, especially when someone in the family has a cold. Wash your hands frequently, particularly before making contact with the infant, especially in a crowded a living space, where there is an increased risk of exposure.

Continued research on both modifiable and non-modifiable risk factors, further prospective validation of risk scoring tools is critical in understanding the factors that most accurately determine RSV-H among 32-35 wGA infants. A stronger understanding of risk factors associated with severe RSV infection will provide policy makers, health care providers, and stakeholders with the necessary evidence based information to assist in decision making regarding RSV prophylaxis, infection control measurements as well as parent education on severe RSV infection.

REFERENCES

1. Hall CB. Respiratory syncytial virus and parainfluenza virus. N Engl J Med 2001 Jun 21;344(25):1917-28.

2. Collins PL, Graham BS. Viral and host factors in human respiratory syncytial virus pathogenesis. J Virol 2008 Mar;82(5):2040-55.

3. Checchia PA, Nalysnyk L, Fernandes AW, Mahadevia PJ, Xu Y, Fahrbach K, Welliver RC S. Mortality and morbidity among infants at high risk for severe respiratory syncytial virus infection receiving prophylaxis with palivizumab: A systematic literature review and meta-analysis. Pediatr Crit Care Med 2011 Sep;12(5):580-8.

4. Langley GF, Anderson LJ. Epidemiology and prevention of respiratory syncytial virus infections among infants and young children. Pediatr Infect Dis J 2011 Jun;30(6):510-7.

5. Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, Anderson LJ. Bronchiolitis-associated hospitalizations among US children, 1980-1996. JAMA 1999 Oct 20;282(15):1440-6.

6. Langley JM, Wang EE, Law BJ, Stephens D, Boucher FD, Dobson S, McDonald J, MacDonald NE, Mitchell I, Robinson JL. Economic evaluation of respiratory syncytial virus infection in canadian children: A pediatric investigators collaborative network on infections in canada (PICNIC) study. J Pediatr 1997 Jul;131(1 Pt 1):113-7.

7. Hall CB, Weinberg GA, Iwane MK, Blumkin AK, Edwards KM, Staat MA, Auinger P, Griffin MR, Poehling KA, Erdman D, et al. The burden of respiratory syncytial virus infection in young children. N Engl J Med 2009 Feb 5;360(6):588-98.

8. Shay DK, Holman RC, Roosevelt GE, Clarke MJ, Anderson LJ. Bronchiolitis-associated mortality and estimates of respiratory syncytial virus-associated deaths among US children, 1979-1997. J Infect Dis 2001 Jan 1;183(1):16-22.

9. Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants. the IMpact-RSV study group Pediatrics 1998 Sep;102(3 Pt 1):531-7.

10. Synagis [Internet] [cited 2011 11/28/2011]. Available from: http://www.abbott.ca/dotca/url/content/en_CA/product/Product_Profile_0199.htm.

11. Law BJ, Langley JM, Allen U, Paes B, Lee DS, Mitchell I, Sampalis J, Walti H, Robinson J, O'Brien K, et al. The pediatric investigators collaborative network on infections in canada study of predictors of hospitalization for respiratory syncytial virus infection for infants born at 33 through 35 completed weeks of gestation. Pediatr Infect Dis J 2004 Sep;23(9):806-14.

12. Prescott WA,Jr, Doloresco F, Brown J, Paladino JA. Cost effectiveness of respiratory syncytial virus prophylaxis: A critical and systematic review. Pharmacoeconomics 2010;28(4):279-93.

13. Roeckl-Wiedmann I, Liese JG, Grill E, Fischer B, Carr D, Belohradsky BH. Economic evaluation of possible prevention of RSV-related hospitalizations in premature infants in germany. Eur J Pediatr 2003 Apr;162(4):237-44.

14. Figueras Aloy J, Carbonell Estrany X, Comite de Estandares de la Sociedad Espanola de Neonatologia. Recommendations for the use of palivizumab in the prevention of respiratory syncytial virus infection in late preterm infants (32(1) to 35(0) weeks of gestation). An Pediatr (Barc) 2010 Aug;73(2):98.e1,98.e4.

15. IWK Health Centre: About Palivizumab (Synagis®) [Internet] [cited 2012 1/4/2012]. Available from: http://www.iwk.nshealth.ca/index.cfm?objectid=86C68F8D-0A4E-E46D-9AA7498758B08B00.

16. Medici MC, Arcangeletti MC, Rossi GA, Lanari M, Merolla R, Paparatti UD, Chezzi C, Osservatorio VRS Study Group. Four year incidence of respiratory syncytial virus infection in infants and young children referred to emergency departments for lower respiratory tract diseases in italy: The "osservatorio VRS" study (2000-2004). New Microbiol 2006 Jan;29(1):35-43.

17. Nielsen HE, Siersma V, Andersen S, Gahrn-Hansen B, Mordhorst CH, Norgaard-Pedersen B, Roder B, Sorensen TL, Temme R, Vestergaard BF. Respiratory syncytial virus infection--risk factors for hospital admission: A case-control study. Acta Paediatr 2003 Nov;92(11):1314-21.

18. Colin AA, McEvoy C, Castile RG. Respiratory morbidity and lung function in preterm infants of 32 to 36 weeks' gestational age. Pediatrics 2010 Jul;126(1):115-28.

19. Boyce TG, Mellen BG, Mitchel EF,Jr, Wright PF, Griffin MR. Rates of hospitalization for respiratory syncytial virus infection among children in medicaid. J Pediatr 2000 Dec;137(6):865-70.

20. Figueras-Aloy J, Carbonell-Estrany X, Quero J, IRIS Study Group. Case-control study of the risk factors linked to respiratory syncytial virus infection requiring hospitalization in premature infants born at a gestational age of 33-35 weeks in spain. Pediatr Infect Dis J 2004 Sep;23(9):815-20.

21. Figueras-Aloy J, Carbonell-Estrany X, Quero-Jimenez J, Fernandez-Colomer B, Guzman-Cabanas J, Echaniz-Urcelay I, Domenech-Martinez E, IRIS Study Group. FLIP-2 study: Risk factors linked to respiratory syncytial virus infection requiring hospitalization in premature infants born in spain at a gestational age of 32 to 35 weeks. Pediatr Infect Dis J 2008 Sep;27(9):788-93.

22. Rossi GA, Medici MC, Arcangeletti MC, Lanari M, Merolla R, Paparatti UD, Silvestri M, Pistorio A, Chezzi C, Osservatorio RSV Study Group. Risk factors for severe RSV-induced lower respiratory tract infection over four consecutive epidemics. Eur J Pediatr 2007 Dec;166(12):1267-72.

23. Holberg CJ, Wright AL, Martinez FD, Ray CG, Taussig LM, Lebowitz MD. Risk factors for respiratory syncytial virus-associated lower respiratory illnesses in the first year of life. Am J Epidemiol 1991 Jun 1;133(11):1135-51.

24. Gooch KL, Notario GF, Schulz G, Gudkov KM, Buesch K, Khong H, Campbell A. Comparison of risk factors between preterm and term infants hospitalized for severe respiratory syncytial virus in the russian federation. Int J Womens Health 2011;3:133-8.

25. Wang EE, Law BJ, Stephens D. Pediatric investigators collaborative network on infections in canada (PICNIC) prospective study of risk factors and outcomes in patients hospitalized with respiratory syncytial viral lower respiratory tract infection. J Pediatr 1995 Feb;126(2):212-9.

26. Chatzimichael A, Tsalkidis A, Cassimos D, Gardikis S, Tripsianis G, Deftereos S, Ktenidou-Kartali S, Tsanakas I. The role of breastfeeding and passive smoking on the development of severe bronchiolitis in infants. Minerva Pediatr 2007 Jun;59(3):199-206.

27. Bachrach VR, Schwarz E, Bachrach LR. Breastfeeding and the risk of hospitalization for respiratory disease in infancy: A meta-analysis. Arch Pediatr Adolesc Med 2003 Mar;157(3):237-43.

28. Tarrant M, Kwok MK, Lam TH, Leung GM, Schooling CM. Breast-feeding and childhood hospitalizations for infections. Epidemiology 2010 Nov;21(6):847-54.

29. Bulkow LR, Singleton RJ, Karron RA, Harrison LH, Alaska RSV Study Group. Risk factors for severe respiratory syncytial virus infection among alaska native children. Pediatrics 2002 Feb;109(2):210-6.

30. Karron RA, Singleton RJ, Bulkow L, Parkinson A, Kruse D, DeSmet I, Indorf C, Petersen KM, Leombruno D, Hurlburt D, et al. Severe respiratory syncytial virus disease in alaska native children. RSV alaska study group. J Infect Dis 1999 Jul;180(1):41-9.

31. Lowther SA, Shay DK, Holman RC, Clarke MJ, Kaufman SF, Anderson LJ. Bronchiolitis-associated hospitalizations among american indian and alaska native children. Pediatr Infect Dis J 2000 Jan;19(1):11-7.

32. Glezen WP, Paredes A, Allison JE, Taber LH, Frank AL. Risk of respiratory syncytial virus infection for infants from low-income families in relationship to age, sex, ethnic group, and maternal antibody level. J Pediatr 1981 May;98(5):708-15.

33. Hand IL, Noble L, Geiss D, Shotkin A. Respiratory syncytial virus immunoprophylaxis in an urban population: A comparison of delivery strategies and outcomes. Pediatr Infect Dis J 2008 Feb;27(2):175-6.

34. Simoes EA, King SJ, Lehr MV, Groothuis JR. Preterm twins and triplets. A high-risk group for severe respiratory syncytial virus infection Am J Dis Child 1993 Mar;147(3):303-6.

35. Respiratory Syncytial Virus -- 2009 (1): 560 -- Red Book [Internet] [cited 2011 11/28/2011]. Available from: http://aapredbook.aappublications.org.ezproxy.library.dal.ca/cgi/content/full/2009/1/3.110?cookietest=yes.

36. Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, O'Brien KL, Roca A, Wright PF, Bruce N, et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: A systematic review and meta-analysis. Lancet 2010 May 1;375(9725):1545-55.

37. Sampalis JS. Morbidity and mortality after RSV-associated hospitalizations among premature canadian infants. J Pediatr 2003 Nov;143(5 Suppl):S150-6.

38. Preventing respiratory syncytial virus infections | Position statements and practice points | Canadian Paediatric Society [Internet] [cited 2012 9/12/2012]. Available from: http://www.cps.ca/en/documents/position/preventing-rsv.

39. CDC Data & Statistics | Feature: Respiratory Syncytial Virus (RSV) Season, U.S., 2010-2011 [Internet] [cited 2011 11/28/2011]. Available from: http://www.cdc.gov/Features/dsRSV/.

40. Sigurs N. Epidemiologic and clinical evidence of a respiratory syncytial virus-reactive airway disease link. Am J Respir Crit Care Med 2001 Mar;163(3 Pt 2):S2-6.

41. Glezen WP, Taber LH, Frank AL, Kasel JA. Risk of primary infection and reinfection with respiratory syncytial virus. Am J Dis Child 1986 June 1;140(6):543-6.

42. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, Fukuda K. Mortality associated with influenza and respiratory syncytial virus in the united states JAMA 2003 Jan 8;289(2):179-86.

43. Carpenter TC, Stenmark KR. Predisposition of infants with chronic lung disease to respiratory syncytial virus-induced respiratory failure: A vascular hypothesis Pediatr Infect Dis J 2004 Jan;23(1 Suppl):S33-40.

44. Buckingham SC, Quasney MW, Bush AJ, DeVincenzo JP. Respiratory syncytial virus infections in the pediatric intensive care unit: Clinical characteristics and risk factors for adverse outcomes. Pediatr Crit Care Med 2001 Oct;2(4):318-23.

45. Meert K, Heidemann S, Lieh-Lai M, Sarnaik AP. Clinical characteristics of respiratory syncytial virus infections in healthy versus previously compromised host. Pediatr Pulmonol 1989;7(3):167-70.

46. Congenital heart disease - PubMed Health [Internet] [cited 2012 1/5/2012]. Available from: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0002103/.

47. Duppenthaler A, Ammann RA, Gorgievski-Hrisoho M, Pfammatter JP, Aebi C. Low incidence of respiratory syncytial virus hospitalisations in haemodynamically significant congenital heart disease. Arch Dis Child 2004 Oct;89(10):961-5.

48. Wang EE, Law BJ, Robinson JL, Dobson S, al Jumaah S, Stephens D, Boucher FD, McDonald J, Mitchell I, MacDonald NE. PICNIC (pediatric investigators collaborative network on infections in canada) study of the role of age and respiratory syncytial virus neutralizing antibody on respiratory syncytial virus illness in patients with underlying heart or lung disease. Pediatrics 1997 Mar;99(3):E9.

49. Sommer C, Resch B, Simoes EA. Risk factors for severe respiratory syncytial virus lower respiratory tract infection Open Microbiol J 2011;5:144-54.

50. Friedrich L, Stein RT, Pitrez PM, Corso AL, Jones MH. Reduced lung function in healthy preterm infants in the first months of life. Am J Respir Crit Care Med 2006 Feb 15;173(4):442-7.

51. Friedrich L, Pitrez PM, Stein RT, Goldani M, Tepper R, Jones MH. Growth rate of lung function in healthy preterm infants. Am J Respir Crit Care Med 2007 Dec 15;176(12):1269-73.

52. Carbonell-Estrany X, Bont L, Doering G, Gouyon JB, Lanari M. Clinical relevance of prevention of respiratory syncytial virus lower respiratory tract infection in preterm infants born between 33 and 35 weeks gestational age. Eur J Clin Microbiol Infect Dis 2008 Oct;27(10):891-9.

53. Carbonell-Estrany X, Figueras-Aloy J, Law BJ, Infeccion Respiratoria Infantil por Virus Respiratorio Sincitial Study Group, Pediatric Investigators Collaborative Network on Infections in Canada Study Group. Identifying risk factors for severe respiratory syncytial virus among infants born after 33 through 35 completed weeks of gestation: Different methodologies yield consistent findings. Pediatr Infect Dis J 2004 Nov;23(11 Suppl):S193-201.

54. Sampalis JS, Langley J, Carbonell-Estrany X, Paes B, O'Brien K, Allen U, Mitchell I, Aloy JF, Pedraz C, Michaliszyn AF. Development and validation of a risk scoring tool to predict respiratory syncytial virus hospitalization in premature infants born at 33 through 35 completed weeks of gestation. Med Decis Making 2008 Jul-Aug;28(4):471-80.

55. Joffe S, Escobar GJ, Black SB, Armstrong MA, Lieu TA. Rehospitalization for respiratory syncytial virus among premature infants. Pediatrics 1999 Oct;104(4 Pt 1):894-9.

56. Liese JG, Grill E, Fischer B, Roeckl-Wiedmann I, Carr D, Belohradsky BH, Munich RSV Study Group. Incidence and risk factors of respiratory syncytial virus-related hospitalizations in premature infants in germany. Eur J Pediatr 2003 Apr;162(4):230-6.

57. Medici MC, Arcangeletti MC, Merolla R, Chezzi C, Osservatorio VRS Study Group. Incidence of respiratory syncytial virus infection in infants and young children referred to the emergency departments for lower respiratory tract diseases in italy. Acta Biomed 2004 Apr;75(1):26-33.

58. Georgescu G, Chemaly RF. Palivizumab: Where to from here? Expert Opin Biol Ther 2009 Jan;9(1):139-47.

59. Hurwitz JL. Respiratory syncytial virus vaccine development. Expert Rev Vaccines 2011 Oct;10(10):1415-33.

60. Committee on Infectious Diseases. From the american academy of pediatrics: Policy statements--modified recommendations for use of palivizumab for prevention of respiratory syncytial virus infections. Pediatrics 2009 Dec;124(6):1694-701.

61. ID11-05.pdf (application/pdf Object) [Internet] [cited 2011 11/29/2011]. Available from: http://www.cps.ca/english/statements/ID/ID11-05.pdf.

62. Langley JM, LeBlanc JC, Wang EE, Law BJ, MacDonald NE, Mitchell I, Stephens D, McDonald J, Boucher FD, Dobson S. Nosocomial respiratory syncytial virus infection in canadian pediatric hospitals: A pediatric investigators collaborative network on infections in canada study. Pediatrics 1997 Dec;100(6):943-6.

63. Simoes EA, Carbonell-Estrany X, Fullarton JR, Rossi GA, Barberi I, Lanari M, European RSV Risk Factor Study Group. European risk factors' model to predict hospitalization of premature infants born 33-35 weeks' gestational age with respiratory syncytial virus: Validation with italian data. J Matern Fetal Neonatal Med 2011 Jan;24(1):152-7.

64. Simoes EA, Carbonell-Estrany X, Fullarton JR, Liese JG, Figueras-Aloy J, Doering G, Guzman J, European RSV Risk Factor Study Group. A predictive model for respiratory syncytial virus (RSV) hospitalisation of premature infants born at 33-35 weeks of gestational age, based on data from the spanish FLIP study. Respir Res 2008 Dec 8;9:78.

65. Carbonell-Estrany X, Quero J, Bustos G, Cotero A, Domenech E, Figueras-Aloy J, Fraga JM, Garcia LG, Garcia-Alix A, Del Rio MG, et al. Rehospitalization because of respiratory syncytial virus infection in premature infants younger than 33 weeks of gestation: A prospective study. IRIS study group. Pediatr Infect Dis J 2000 Jul;19(7):592-7.

66. Stevens TP, Sinkin RA, Hall CB, Maniscalco WM, McConnochie KM. Respiratory syncytial virus and premature infants born at 32 weeks' gestation or earlier: Hospitalization and economic implications of prophylaxis. Arch Pediatr Adolesc Med 2000 Jan;154(1):55-61.

67. Carbonell-Estrany X, Quero J, IRIS Study Group. Hospitalization rates for respiratory syncytial virus infection in premature infants born during two consecutive seasons. Pediatr Infect Dis J 2001 Sep;20(9):874-9.

68. Broughton S, Roberts A, Fox G, Pollina E, Zuckerman M, Chaudhry S, Greenough A. Prospective study of healthcare utilisation and respiratory morbidity due to RSV infection in prematurely born infants. Thorax 2005 Dec;60(12):1039-44.

69. Difranza JR, Masaquel A, Barrett AM, Colosia AD, Mahadevia PJ. Systematic literature review assessing tobacco smoke exposure as a risk factor for serious respiratory syncytial virus disease among infants and young children BMC Pediatr 2012 Jun 21;12:81.

70. Committee on Infectious Diseases. From the american academy of pediatrics: Policy statements--modified recommendations for use of palivizumab for prevention of respiratory syncytial virus infections Pediatrics 2009 Dec;124(6):1694-701.

71. pnr_comp_doc_july2007.pdf (application/pdf Object) [Internet] [cited 2012 2/1/2012]. Available from: http://rcp.nshealth.ca/sites/default/files/clinical-practice-guidelines/pnr_comp_doc_july2007.pdf.

72. Reproductive Care Program of Nova Scotia | [Internet] [cited 2012 8/29/2012]. Available from: http://rcp.nshealth.ca/.

73. Groothuis JR, Fryzek JP, Makari D, Steffey D, Martone WJ. Respiratory syncytial virus hospitalization trends in infants with chronic lung disease of infancy, 1998-2008 Clin Epidemiol 2011;3:245-50.

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

75. Simoes EA. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease J Pediatr 2003 Nov;143(5 Suppl):S118-26.

76. Law B, Macdonald N, Langley J, Mitchell I, Stephens D, Wang E, Robinson J, Boucher F, McDonald J, Dobson S. Severe respiratory syncytial virus infection among otherwise healthy prematurely born infants: What are we trying to prevent? Paediatr Child Health 1998 Nov;3(6):402-4.

77. Okiro EA, Ngama M, Bett A, Cane PA, Medley GF, James Nokes D. Factors associated with increased risk of progression to respiratory syncytial virus-associated pneumonia in young kenyan children Trop Med Int Health 2008 Jul;13(7):914-26.

78. Rietveld E, Vergouwe Y, Steyerberg EW, Huysman MW, de Groot R, Moll HA, RSV Study Group Southwest Netherlands. Hospitalization for respiratory syncytial virus infection in young children: Development of a clinical prediction rule Pediatr Infect Dis J 2006 Mar;25(3):201-7.

79. Kramer MS, Platt RW, Wen SW, Joseph KS, Allen A, Abrahamowicz M, Blondel B, Breart G, Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance System. A new and improved population-based canadian reference for birth weight for gestational age Pediatrics 2001 Aug;108(2):E35.

80. Martinez FD, Morgan WJ, Wright AL, Holberg CJ, Taussig LM. Diminished lung function as a predisposing factor for wheezing respiratory illness in infants N Engl J Med 1988 Oct 27;319(17):1112-7.

81. Colosia AD, Masaquel A, Hall CB, Barrett AM, Mahadevia PJ, Yogev R. Residential crowding and severe respiratory syncytial virus disease among infants and young children: A systematic literature review BMC Infect Dis 2012 Apr 20;12(1):95.

82. APGAR: MedlinePlus Medical Encyclopedia [Internet] [cited 2012 9/10/2012]. Available from: http://www.nlm.nih.gov/medlineplus/ency/article/003402.htm.

83. Simoes EA. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease. J Pediatr 2003 Nov;143(5 Suppl):S118-26.

84. Stensballe LG, Ravn H, Kristensen K, Meakins T, Aaby P, Simoes EA. Seasonal variation of maternally derived respiratory syncytial virus antibodies and association with infant hospitalizations for respiratory syncytial virus J Pediatr 2009 Feb;154(2):296-8.

85. Data Documentation | Reproductive Care Program of Nova Scotia [Internet] [cited 2011 12/1/2011]. Available from: http://rcp.nshealth.ca/atlee-database/data-documentation.

86. Discharge Abstract Database [Internet] [cited 2011 12/1/2011]. Available from: http://www.cihi.ca/cihi-ext-portal/internet/en/document/types+of+care/hospital+care/acute+care/services_dad.

86. Discharge Abstract Database [Internet] [cited 2011 12/1/2011]. Available from: http://www.cihi.ca/cihi-ext-portal/internet/en/document/types+of+care/hospital+care/acute+care/services_dad.

87. CIHI - ICD-10 Coding Structure [Internet] [cited 2012 8/13/2012]. Available from: http://www.cihi.ca/CIHI-ext-portal/internet/en/document/standards+and+data+submission/standards/classification+and+coding/codingclass_icd10struct

88. Statistics Canada: Canada's national statistical agency [Internet] [cited 2012 11/6/2012]. Available from: http://www.statcan.gc.ca/start-debut-eng.html.

89. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve Am J Epidemiol 2006 Apr 1;163(7):670-5.

90. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: Clinical example of sepsis Intensive Care Med 2003 Jul;29(7):1043-51.91. Sensitivity, Specificity, and Predictive Value - Clinical Methods - NCBI Bookshelf [Internet] [cited 2012 10/18/2012]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK383/.

92. Deeks JJ, Altman DG. Diagnostic tests 4: Likelihood ratios BMJ 2004 Jul 17;329(7458):168-9.

93. McGee S. Simplifying likelihood ratios Journal of General Internal Medicine 2002;17(8):647 <last_page> 650.

94. ICD - ICD-9-CM - International Classification of Diseases, Ninth Revision, Clinical Modification [Internet] [cited 2012 8/13/2012]. Available from: http://www.cdc.gov/nchs/icd/icd9cm.htm.

95. ICD - ICD-10-CM - International Classification of Diseases, Tenth Revision, Clinical Modification [Internet] [cited 2012 8/13/2012]. Available from: http://www.cdc.gov/nchs/icd/icd10cm.htm.

96. Tregoning JS, Schwarze J. Respiratory viral infections in infants: Causes, clinical symptoms, virology, and immunology Clin Microbiol Rev 2010 Jan;23(1):74-98.

97. rdc0901e.pdf (application/pdf Object) [Internet] [cited 2012 8/8/2012]. Available from: http://www.phac-aspc.gc.ca/publicat/rdc-mrc01/pdf/rdc0901e.pdf.

98. Bradley JP, Bacharier LB, Bonfiglio J, Schechtman KB, Strunk R, Storch G, Castro M. Severity of respiratory syncytial virus bronchiolitis is affected by cigarette smoke exposure and atopy Pediatrics 2005 Jan;115(1):e7-14.

99. Harrod KS, Jaramillo RJ, Rosenberger CL, Wang SZ, Berger JA, McDonald JD, Reed MD. Increased susceptibility to RSV infection by exposure to inhaled diesel engine emissions Am J Respir Cell Mol Biol 2003 Apr;28(4):451-63.

100. Linnet K, Bossuyt PM, Moons KG, Reitsma JB. Quantifying the accuracy of a diagnostic test or marker. Clin Chem 2012 Sep;58(9):1292-301.

101. Stensballe LG, Fullarton JR, Carbonell-Estrany X, Simoes EA. Population based external validation of a european predictive model for respiratory syncytial virus hospitalization of premature infants born 33 to 35 weeks of gestational age Pediatr Infect Dis J 2010 Apr;29(4):374-6.

102. Carbonell-Estrany X, Simoes EA, Fullarton JR, Ferdynus C, Gouyon JB, European RSV Risk Factor Study Group. Validation of a model to predict hospitalization due to RSV of infants born at 33-35 weeks' gestation. J Perinat Med 2010 Jul;38(4):411-7.

103. Canadian Perinatal Health Report - 2008 edition - Public Health Agency of Canada [Internet] [cited 2012 10/23/2012]. Available from: http://www.phac-aspc.gc.ca/publicat/2008/cphr-rspc/index-eng.php.

104. Geronimus AT, Bound J. Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. Am J Epidemiol 1998 Sep 1;148(5):475-86.

105. Windham GC, Swan SH, Fenster L. Parental cigarette smoking and the risk of spontaneous abortion Am J Epidemiol 1992 Jun 15;135(12):1394-403.

106. Tobacco use among U.S. racial/ethnic minority groups--african americans, american indians and alaska natives, asian americans and pacific islanders, hispanics. A report of the surgeon general. executive summary MMWR Recomm Rep 1998 Oct 9;47(RR-18):v,xv, 1-16.

107. Shea AK, Steiner M. Cigarette smoking during pregnancy Nicotine Tob Res 2008 Feb;10(2):267-78.

108. Myers D, Lee SW. Immigration cohorts and residential overcrowding in southern california Demography 1996 Feb;33(1):51-65.

109. Haustein KO. Cigarette smoking, nicotine and pregnancy Int J Clin Pharmacol Ther 1999 Sep;37(9):417-27.

110. DiFranza JR, Lew RA. Effect of maternal cigarette smoking on pregnancy complications and sudden infant death syndrome J Fam Pract 1995 Apr;40(4):385-94.

111. Kovesi T, Gilbert NL, Stocco C, Fugler D, Dales RE, Guay M, Miller JD. Indoor air quality and the risk of lower respiratory tract infections in young canadian inuit children. CMAJ 2007 Jul 17;177(2):155-60.

112. Young M, Kandola K, Mitchell R, Leamon A. Hospital admission rates for lower respiratory tract infections in infants in the northwest territories and the kitikmeot region of nunavut between 2000 and 2004. Paediatr Child Health 2007 Sep;12(7):563-6.

113. Simoes EA. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease. J Pediatr 2003 Nov;143(5 Suppl):S118-26.

114. Simoes EA. Maternal smoking, asthma, and bronchiolitis: Clear-cut association or equivocal evidence? Pediatrics 2007 Jun;119(6):1210-2.

115. Preventing respiratory syncytial virus infections | Position statements and practice points | Canadian Paediatric Society [Internet] [cited 2012 8/17/2012]. Available from: http://www.cps.ca/en/documents/position/preventing-rsv.

116. Figueras-Aloy J, Quero-Jimenez J, Fernandez-Colomer B, Guzman-Cabanas J, Echaniz-Urcelay I, Domenech-Martinez E, Grupo IRIS. Usefulness of different risk factor associations in predicting admissions due to respiratory syncytial virus in premature newborns of 32 to 35 weeks gestation in spain. An Pediatr (Barc) 2009 Jul;71(1):47-53.

117. Banerji A, Greenberg D, White LF, Macdonald WA, Saxton A, Thomas E, Sage D, Mamdani M, Lanctot KL, Mahony JB, et al. Risk factors and viruses associated with hospitalization due to lower respiratory tract infections in canadian inuit children : A case-control study. Pediatr Infect Dis J 2009 Aug;28(8):697-701.

APPENDIX A: Risk factors presented in the literature and predictive variables available through the Atlee database as either proxies or direct measures

Risk factors (Literature)	Direct measure or proxy	Predictive variable available in the Atlee
Month of birth [11, 64, 83]	Direct measure	Date of birth (MM/DD/YYYY)
Gestational age (GA)	Direct measure	GA of infant at birth; 32 weeks; 33 weeks; 34 weeks; and 35 weeks
Chronological age (CA)[22, 23, 64, 116]	Direct measure	Date of birth (MM/DD/YYYY) used as a calculation of subject's CA on December 1st of the same year of study.
Preschool aged siblings [11, 116]	Proxy	Gravida- previous births of children ≤ 5 and the assumption made that they are still in the household
Birth weight (grams) [78]	Direct measure	Birth weight in grams (g)
Small for gestational age [11]	Direct measure	Category for birth weight for GA according to Kramer's criteria for Canadian premature infant's based on GA and sex of the infant (<3rd -97th percentile)) [79]
Infant sex [11, 23, 32, 56, 64, 83, 116]	Direct measure	Male or female
Passive household smoke exposure [11, 69, 83, 116, 117]	Direct Measure	Mother's smoking status at delivery date used as a measurement for passive household smoke exposure
Household crowding/number of siblings[11, 22-24, 81, 83, 116, 117]	Proxy	Household crowding (≥ 5 individuals per household) was inferred from the number of previous deliveries in proceeding 18 years (gravida) and marital status (single vs. married) at the time the baby was born
Low socioeconomic status (SES)/maternal education [32, 83]	Proxy	SES-Mother's postal code linked to Canadian census used as an approximation for neighbourhood income quintile (performed by the RCP)
Breast feeding[23, 24, 27, 64, 83, 116, 117]	Direct Measure	Mother's breast feeding status at birth
Apgar score at 1 and 5 minutes	Direct measure	Score between 0-10 both at 1 and 5 minutes
Mode of delivery	Direct measure	C-section or Vaginal
Maternal age [24]	Direct measure	Mother's age at time of delivery
Multiple births [24]	Direct measure	Number of births; singleton vs. multiples (twins/triplets)

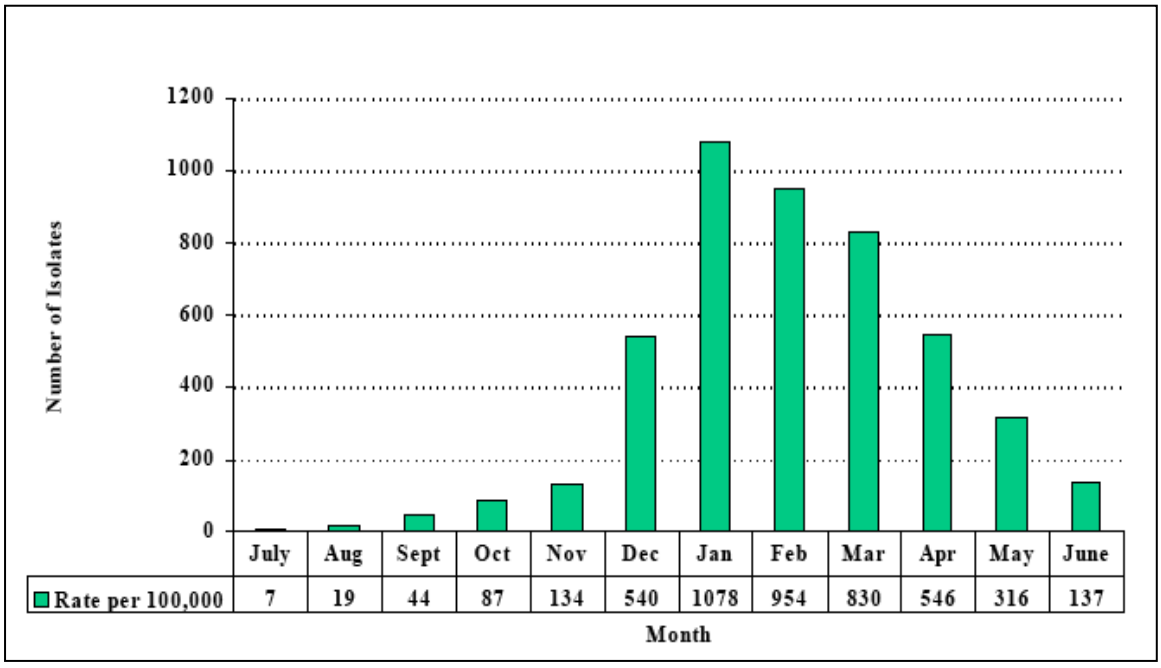
Risk factors (Literature)	Direct measure or proxy	Predictive variable available in the Atlee
Premature preterm rupture of membranes (PPRoM)	Direct measure	PPRoM during current pregnancy
Antenatal steroids	Direct measure	The use of antenatal steroid during pregnancy
Daycare Attendence (subject and or siblings) [11, 22, 23, 56, 83]	**NA**	**NA**
Family history of allergies/eczema/ respiratory disease [11, 64, 116]	**NA**	**NA**
Low maternal RSV antibody levels [23, 32, 84]	**NA**	**NA**
Race/ethnicity [30, 31]	**NA**	**Not available in Atlee during study period**

APPENDIX B: Detailed description of variables included for defining study eligibility criteria, predictive variables for follow up analysis and outcome of interest

Variable	Database	Details from the Database
Inclusion Variable		
Infant date of birth (YYYY/MM/DD)	Atlee	Live birth between July 1st, 1998 and June 31st, 2008 in hospital setting
Mother's postal code	Atlee	Mother's postal code as a measurement of the infant's residency in NS
Gestational age	Atlee	32 weeks 0 days to 35 weeks 6 days
Exclusion Variable		
RSV-Palivizumab immunoprophylaxis	NS Blood Coordinating Program database of Pz Recipients	Subjects will be excluded if the infant has previously been administered RSV immunoprophylaxis (Pz)
CLD/BPD	Atlee	Infants with underlying medical condition (CLD)/(BPD) who automatically qualify for the prophylaxis: Variables include date of birth, sex, and the infant's eligibility criteria
CHD	IWK Cardiology Database	Infants with underlying medical condition (CHD): Data elements of direct measurement include date of birth, sex, cardiac diagnostic fields, Pz eligibility status (since 2003), and corrective surgery date.
Outcome after birth discharge	Atlee	-Early neonatal death (0-7 days) -Late neonatal death (7-38 days)
Predictive Variables		
Month of birth	Atlee	Date of birth (MM/DD/YYYY)
Gestational age (GA)	Atlee	GA of infant at birth; 32 weeks; 33 weeks; 34 weeks; and 35 weeks
Chronological age (CA)	Atlee	Date of birth (MM/DD/YYYY) used as a calculation of subject's CA on December 1st of the same year of study.
Preschool aged siblings	Atlee	Gravida- previous births of children ≤ 5 and the assumption made that they are still in the household
Birth weight (g)	Atlee	Birth weight in grams (g)
Small for GA		Category for birth weight for GA according to Kramer's criteria for Canadian premature infant's based on GA and sex of the infant [79]
Infant sex	Atlee	Male or female

Variable	Database	Details from the Database
Passive household smoke exposure	Atlee	Mother's smoking status at delivery date used as a measurement for passive household smoke exposure
Household crowding/number of siblings	Atlee	Household crowding (≥ 5 individuals per household) was inferred from the number of previous deliveries in proceeding 18 years (gravida) and marital status (single vs. married) at the time the baby was born
Low socioeconomic status (SES)	Atlee/ Canadian census	SES-Mother's postal code linked to Canadian census used as an approximation for neighbourhood income quintile (performed by the RCP)
Breast feeding	Atlee	Mother's breast feeding status at birth
Apgar score at 1 and 5 minutes	Atlee	Score between 0-10 both at 1 and 5 minutes
Mode of delivery	Atlee	C-section or Vaginal
Maternal age	Atlee	Mother's age at time of delivery
Multiple births	Atlee	Number of births; singleton vs. multiples (twins/triplets)
Premature preterm rupture of membranes (PPRoM)	Atlee	PPRoM during current pregnancy
Antenatal steroids	Atlee	The use of antenatal steroid during pregnancy
Outcome Variables		
RSV-associated severe LRTI hospitalization (RSV-H)	CIHI	Infant's primary hospitalization attributable to RSV-associated severe LRTI (ICD-CM codes)

APPENDIX C: Number of RSV positive laboratory isolates per month in Canada (1997-1998)



Source: Notifiable Disease Reporting System, Division of Disease Surveillance, Centre for Infectious Disease Prevention and Control, Health Canada [98].
Note: RSV activity in the Northern Hemisphere usually starts in the late fall, peaks in early winter and tapers off in the late spring.