A Systematic Review of Predictive Modeling for Bronchiolitis

Gang Luo (corresponding author)
Department of Biomedical Informatics, University of Utah
Suite 140, 421 Wakara Way, Salt Lake City, UT 84108, USA
gang.luo@utah.edu
Phone: 1-801-213-3565

Flory L. Nkoy, Per H. Gesteland, Tiffany S. Glasgow, Bryan L. Stone
Department of Pediatrics, University of Utah
100 N Mario Capecchi Drive, Salt Lake City, UT 84113, USA
flory.nkoy@hsc.utah.edu, per.gesteland@hsc.utah.edu, tiffany.glasgow@hsc.utah.edu, bryan.stone@hsc.utah.edu

Abstract

Purpose: Bronchiolitis is the most common cause of illness leading to hospitalization in young children. At present, many bronchiolitis management decisions are made subjectively, leading to significant practice variation among hospitals and physicians caring for children with bronchiolitis. To standardize care for bronchiolitis, researchers have proposed various models to predict the disease course to help determine a proper management plan. This paper reviews the existing state of the art of predictive modeling for bronchiolitis. Predictive modeling for respiratory syncytial virus (RSV) infection is covered whenever appropriate, as RSV accounts for about 70% of bronchiolitis cases.

Methods: A systematic review was conducted through a PubMed search up to April 25, 2014. The literature on predictive modeling for bronchiolitis was retrieved using a comprehensive search query, which was developed through an iterative process. Search results were limited to human subjects, the English language, and children (birth-18 years).

Results: The literature search returned 2,312 references in total. After manual review, 168 of these references were determined to be relevant and are discussed in this paper. We identify several limitations and open problems in predictive modeling for bronchiolitis, and provide some preliminary thoughts on how to address them, with the hope to stimulate future research in this domain.

Conclusions: Many problems remain open in predictive modeling for bronchiolitis. Future studies will need to address them to achieve optimal predictive models.

Keywords: Bronchiolitis, predictive modeling, machine learning, respiratory syncytial virus
1. Introduction

Bronchiolitis is inflammation of the smallest air passages in the lungs (bronchioles) and primarily a disease of children younger than two years. Bronchiolitis is the leading cause of and accounts for 16% of all infant hospitalizations [1-3]. 10% of children are affected by bronchiolitis in their first year of life [4]. By age two, more than one third of children have experienced bronchiolitis [5]. In the U.S., bronchiolitis incurs an annual total hospitalization cost of $543 million [6], and consumes significant emergency department (ED) and hospital resources. Bronchiolitis leads to approximately 238 outpatient visits, 71 hospitalizations, and 77 ED visits per 1,000 infant years [7]. About 30% of infants with bronchiolitis evaluated in pediatric EDs are hospitalized [2]. Overall, about 10% of children with bronchiolitis are hospitalized [5]. Between 2% and 6% of all children with bronchiolitis require care in an intensive care unit (ICU) [8].

A variety of therapies, such as bronchodilators, are used in bronchiolitis with little supporting evidence [9], and minimal consensus on their use other than recommending that clinicians individualize care based on course and severity. Perhaps the only exception is the recommendation to use infection control procedures, but even the extent of this intervention is unclear beyond using hand decontamination.

In evaluating and treating bronchiolitis, a key step is an attempt to anticipate the disease course to guide the appropriate management setting and intensity [8]. At present, many bronchiolitis management decisions are made subjectively [2, 10]. This leads to significant practice variation, as is reflected in variable admission rates and use of specific therapies among different hospitals and physicians [1, 4, 10-20]. Observed practice variation is not explained by differences in patient severity and has little impact on outcomes, but has a significant impact on healthcare resource usage [17]. Excessive hospital admission leads to overuse of inpatient resources, exposes patients to unnecessary iatrogenic risks and other infectious diseases in the hospital, and unnecessarily exposes other hospitalized children to these patients’ infectious respiratory pathogens [12, 18, 21]. One study [22] suggests that up to 10% of infants with bronchiolitis experience adverse events during their hospital stay. Alternatively, patients not properly admitted risk inadequate treatment and medical deterioration including death [12]. Thus, it is desirable to develop methodologies to standardize bronchiolitis care, which can help reduce healthcare cost and improve patient safety and outcome [21, 23, 24].

One way to standardize care for bronchiolitis is to develop and use clinical practice guidelines [8, 9]. With proper implementation, clinical practice guidelines for bronchiolitis can reduce healthcare resource usage by up to 77% without negatively impacting clinical outcomes or patient family’s satisfaction [23, 25-28]. However, due to an insufficient level of detail and limited amounts of evidence, existing clinical practice guidelines provide guidance for a limited number of patients
and still rely heavily on individualized clinician judgment. More detailed guidelines are difficult to generalize because they cannot answer the many combinations of patient and illness characteristics, such as comorbidities.

Another way to standardize care for bronchiolitis is to develop predictive models [21, 29-37] and use them to help direct an optimal disease management plan. By using a data-driven approach to summarize useful information accumulated in clinical and administrative data sets, predictive models [38] can manage the level of individualized detail inherent in a clinical setting, complement clinical practice guidelines, and overcome their limitations. Predictive models are often integrated into computerized decision support tools [273]. These tools can support clinicians’ provisional judgment, or lead clinicians to question and reconsider that judgment [31]. This is particularly useful for inexperienced junior physicians and physicians who see children relatively infrequently. In general, human experts usually make better decisions when they are provided with predictive models’ computational results [39, page 6].

In this paper, we present an overview of existing predictive models for clinical management of bronchiolitis and disease outcomes as well as their limitations. We identify several knowledge gaps and opportunities for improving predictive modeling for bronchiolitis, which can help direct the proper care setting and management for children with bronchiolitis. We discuss how to use machine learning techniques to address some of the gaps and limitations, and hope this paper can stimulate future research on predictive modeling for bronchiolitis. Our paper also covers predictive modeling for respiratory syncytial virus (RSV) infection whenever appropriate, as RSV accounts for about 70% of bronchiolitis cases [40] and has a richer predictive modeling literature base. A list of acronyms used in this paper is provided in the appendix.

2. Methods

The methodology of our study follows the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guideline [272]. The study protocol was designed by and iteratively refined with inputs from all study co-authors.

We conducted a systematic review limited to PubMed by developing a search strategy to retrieve the literature on predictive modeling for bronchiolitis through April 25, 2014. We started from an initial straightforward search query: *(bronchiolitis or RSV) AND prediction*. For each retrieved reference, two independent reviewers (GL and BS) evaluated the title and abstract to determine potential relevancy. Relevancy was judged based on pre-defined inclusion criteria ensuring that the article’s primary focus addressed predictive modeling for some aspect of the clinical management of bronchiolitis/RSV infection and/or disease outcomes. We included articles describing predictors of or risk factors for outcomes regardless of their levels of scientific evidence, in an effort to investigate as many predictors and risk factors as possible to include all factors that might increase the accuracy of machine learning predictive models. If a reference appeared to be potentially relevant, the full text was evaluated
to make a final inclusion decision. We also reviewed the citations in the included articles. If a citation was found to be relevant and missing from the original search results, a keyword phrase was extracted from this citation and added to the search query for expansion to include the relevant citation and other similar articles.

This process was repeated iteratively until a comprehensive search query was developed. Many keyword phrases added to the search query were forms of or synonyms of the words *prediction* and *RSV*. We also found that we could use specific keyword phrases to exclude irrelevant articles and narrow our search results. The final search query used was (*bronchiolitis or RSV or “respiratory syncytial virus” or “respiratory syncytial viral”*) AND (*predict OR predicting OR predicts or predicted OR prediction OR predictive OR predictor OR predictors OR “risk scoring” OR “risk factor” OR “risk factors” OR asthma OR epidemiologic* OR management OR update*) NOT (*bronchiectasis OR malignancy OR “bronchiolitis obliterans” OR transplant*). Search results were limited to human subjects, the English language, and children (birth-18 years). The final literature review included articles meeting the pre-defined inclusion criteria. Disagreements about inclusion of individual articles were addressed by discussion among GL and BS, and if needed a third reviewer (FN). One reviewer (GL) extracted from the included articles information on five dimensions: purpose for making the prediction, patient population, methods used for building predictive models, risk factors and/or predictors identified, and performance of the predictive models. Any issues involving uncertainty were resolved through discussion among GL and BS, and if needed a third reviewer (FN).

3. Results

![Flowchart of the article selection process.](image)

As shown in Fig. 1, the literature search returned 2,312 references in total. 201 appeared to be potentially relevant after review of titles and abstracts, and underwent full-text review. 168 of these were determined to be relevant and are discussed in this paper. The included articles are primarily predictive modeling studies and observational cohort studies regarding identifying predictors and risk factors, and contain only two systematic reviews and one randomized controlled trial illustrating
the gap in high-quality evidence addressing this subject. In this section, we describe the state of the art of the predictive modeling areas of clinical interest for bronchiolitis, and identify open problems in predictive modeling for bronchiolitis. A summary of the primary clinical research topics on predictive modeling for bronchiolitis is provided in Table 1. Our categorization of these topics is based on input from study co-authors who are domain experts in pediatrics (FN, PG, TG, and BS) and articulated the major problems regarding bronchiolitis management in multiple iterations.

**Table 1.** Categorization and current status of the primary clinical research topics on predictive modeling for bronchiolitis.

<table>
<thead>
<tr>
<th>Primary category</th>
<th>Sub-category</th>
<th>Methods for building existing predictive models</th>
<th>Predictors and/or risk factors identified?</th>
<th>Open problems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predict optimal escalation and de-escalation of care setting for bronchiolitis patients</strong></td>
<td>Predict hospital admission from the ED</td>
<td>Logistic regression, neural network ensemble, scoring system, CART</td>
<td>Yes</td>
<td>Predict the true need for hospital admission for bronchiolitis patients in the ED.</td>
</tr>
<tr>
<td></td>
<td>Predict hospital admission in the ED observation unit</td>
<td>None</td>
<td>Yes</td>
<td>Predict immediately before observation unit admission, whether a bronchiolitis patient will eventually have a true need for hospital admission.</td>
</tr>
<tr>
<td></td>
<td>Predict ICU admission</td>
<td>Logistic regression, applying a threshold to a single clinical parameter</td>
<td>Yes</td>
<td>(1) Predict with high accuracy ICU transfer need for a patient admitted to the general inpatient ward with bronchiolitis (or RSV infection). (2) Predict the true need for ICU admission for bronchiolitis patients in the ED.</td>
</tr>
<tr>
<td></td>
<td>Predict optimal disposition in the primary care setting</td>
<td>None</td>
<td>Yes</td>
<td>Predict in the primary care setting, whether a bronchiolitis patient will use acute care for bronchiolitis in the near future.</td>
</tr>
<tr>
<td></td>
<td>Predict safe discharge and unscheduled visit</td>
<td>Logistic regression</td>
<td>No</td>
<td>Predict upon ED discharge, whether a bronchiolitis patient will have an unscheduled visit in the very near future.</td>
</tr>
<tr>
<td><strong>Make predictions related to respiratory support</strong></td>
<td>Predict whether a bronchiolitis patient will develop apnea</td>
<td>Risk criteria</td>
<td>Yes</td>
<td>Predict with high accuracy whether a bronchiolitis patient in the ED, general inpatient ward, or ICU will develop apnea.</td>
</tr>
<tr>
<td></td>
<td>Predict the use of a specific type of NIPPV</td>
<td>None</td>
<td>Yes</td>
<td>Predict whether a bronchiolitis patient in the ED or general inpatient ward will eventually use a specific type of NIPPV.</td>
</tr>
<tr>
<td></td>
<td>Predict NIPPV failure</td>
<td>None</td>
<td>Yes</td>
<td>Predict whether a bronchiolitis patient in the hospital requiring positive pressure ventilation support will fail NIPPV.</td>
</tr>
<tr>
<td></td>
<td>Predict the use of supplemental oxygen</td>
<td>Bronchiolitis severity score</td>
<td>Yes</td>
<td>(1) Predict with high accuracy whether a bronchiolitis patient in the ED or inpatient ward will eventually use supplemental oxygen. (2) Predict the number of days for which a bronchiolitis patient in the ED or inpatient ward will use supplemental oxygen.</td>
</tr>
<tr>
<td></td>
<td>Predict the use of mechanical ventilation</td>
<td>Logistic regression</td>
<td>Yes</td>
<td>Predict with high accuracy the need of mechanical ventilation in bronchiolitis patients.</td>
</tr>
</tbody>
</table>
### Predictions on Bronchiolitis Patients

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Methodology</th>
<th>Success Rate</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict prolonged mechanical ventilation</td>
<td>Logistic regression</td>
<td>Yes</td>
<td>Predict prolonged mechanical ventilation on a patient who is admitted to the ICU with bronchiolitis and requires mechanical ventilation.</td>
</tr>
<tr>
<td>Predict the use of extracorporeal membrane oxygenation</td>
<td>None</td>
<td>Yes</td>
<td>Predict whether a patient hospitalized with RSV bronchiolitis will need extracorporeal membrane oxygenation support.</td>
</tr>
<tr>
<td>Predict extubation failure</td>
<td>None</td>
<td>Yes</td>
<td>Predict extubation failure in bronchiolitis patients in the ICU for all ventilation modes.</td>
</tr>
<tr>
<td>Make predictions about the value of tests and other evaluations on a</td>
<td>Predict serious bacterial infection</td>
<td>None</td>
<td>Predict serious bacterial infection for patients hospitalized with bronchiolitis.</td>
</tr>
<tr>
<td>bronchiolitis patient</td>
<td>Predict an unhelpful chest X-ray</td>
<td>Logistic regression</td>
<td>No</td>
</tr>
<tr>
<td>Identify RSV infection</td>
<td>Logistic regression</td>
<td>Yes</td>
<td>Accurately identify RSV infection and provide reliable prediction for all patient populations suspected of having RSV infection.</td>
</tr>
<tr>
<td>Make predictions regarding the current hospitalization</td>
<td>Predict hospital length of stay</td>
<td>Logistic regression, CART, neural network ensemble</td>
<td>Yes</td>
</tr>
<tr>
<td>Predict hospitalization cost</td>
<td>Linear regression</td>
<td>Yes</td>
<td>Predict with high accuracy hospitalization cost of a patient to be admitted for bronchiolitis.</td>
</tr>
<tr>
<td>Make predictions related to the use of palivizumab</td>
<td>Predict RSV hospitalization</td>
<td>Logistic regression, discriminatory function analysis</td>
<td>Yes</td>
</tr>
<tr>
<td>Predict compliance with palivizumab</td>
<td>None</td>
<td>Yes</td>
<td>Predict compliance with palivizumab.</td>
</tr>
<tr>
<td>Predict community RSV activity</td>
<td>Naive Bayes, autoregressive integrated moving</td>
<td>Yes</td>
<td>Predict community RSV activity with high accuracy.</td>
</tr>
<tr>
<td>Predict whether a bronchiolitis patient will later be diagnosed with asthma</td>
<td>Logistic regression</td>
<td>Yes</td>
<td>Predict with high accuracy whether a bronchiolitis patient will later be diagnosed with asthma.</td>
</tr>
</tbody>
</table>

As explained in Sections 3.9 and 3.10 in detail, almost every existing predictive model for bronchiolitis was developed for a subset of bronchiolitis patients rather than for all bronchiolitis patients. In addition, most of the existing predictive models for bronchiolitis have inadequate accuracy. Moreover, many previous studies only identified various predictors or risk factors for managing bronchiolitis, but did not develop any predictive model for bronchiolitis.

### 3.1 Predicting optimal escalation and de-escalation of care setting for bronchiolitis patients

#### 3.1.1 Predicting hospital admission from the ED

Much existing work on predictive modeling for bronchiolitis is related to predicting hospital admission in the ED setting.

Walsh *et al.* [36] used neural network ensemble to predict whether a bronchiolitis patient in the ED will be admitted or discharged, with an accuracy of 81%. In comparison, the admitting resident physicians made correct decisions 77% of the time, using an attending pediatrician’s judgment and a length of stay longer than one day as the comparator gold standard.
Marlais et al. [31] used a scoring system to predict whether a bronchiolitis patient in the ED will be admitted, with an Area Under the receiver operating characteristic Curve (AUC) of 81%. This study included only children no more than one year old. Destino et al. [41] used the Children’s Hospital of Wisconsin Respiratory Score to make the same prediction, with a low sensitivity of 65%, a low specificity of 65%, and a low AUC of 68%.

Using nasal wash lactate dehydrogenase concentration level, Laham et al. [42] built a logistic regression model to predict whether a bronchiolitis patient in the ED will be admitted. The predictive model achieved an AUC of 87%, a classification accuracy of 80%, a sensitivity of 81%, a specificity of 77%, a positive predictive value of 88%, and a negative predictive value of 66%.

Corneli et al. [43] used Classification And Regression Tree (CART) to predict whether an infant with bronchiolitis in the ED will be admitted, with a low sensitivity of 56% and a specificity of 74%. This study used various criteria to exclude most infants with bronchiolitis in the ED.

For bronchiolitis patients in the ED, Walsh et al. [24] built a logistic regression model to predict the need for hospital admission, which was defined as actual hospital admission, discharge with a subsequent return visit requiring an admission, or clearly inappropriate discharge. This definition does not reflect a true need for admission because some admissions are unnecessary [21]. The developed model achieved a sensitivity of 91%, a specificity of 83%, and a low positive predictive value of 62%.

In addition to the articles constructing predictive models, there are some other articles describing predictors. For bronchiolitis patients in the ED, Al-Shehri et al. [44-48] identified several predictors of hospital admission: prematurity, chronic lung disease (primarily bronchopulmonary dysplasia), atopic dermatitis, pure formula feeding (no breast feeding), passive smoking, age ≤ one year, nasopharyngeal lactate dehydrogenase value, low dew point, enterovirus infection, absence of familial atopy, rhinovirus, and co-infection. Voets et al. [35] identified three predictors of hospital and ICU admission: age < six months, respiratory rate > 45 breaths per minute, and oxygen saturation < 95% at sea level. This study excluded patients with chronic lung disease, prematurity, underlying cardiac disease, or neurological diseases.

For previously healthy infants with RSV infection, Somech et al. [49] identified several predictors of hospital admission: fever, SaO2 level, abnormal chest auscultation findings, abnormal chest X-ray, and the clinical severity score suggested by Wang et al. [50]. Parker et al. [2] identified several predictors of major medical interventions in bronchiolitis patients, including oxygen administration for ≥ 30 minutes for saturation < 90% in room air, intravenous fluid bolus of ≥ 20ml/kg, any treatment
for apnea, and ICU admission. This study excluded infants with bronchiolitis who were previously unhealthy and provided no predictive model.

Some hospital admissions are unnecessary [21], and some ED discharges to home should have been admitted. It is important to identify the true need for hospital admission rather than only the occurrence of hospital admission. Although some of the work mentioned above achieved reasonably good prediction accuracies, none of the work mentioned above has predicted the true need for hospital admission for bronchiolitis patients in the ED. This remains an open problem. A true need for hospital admission is defined as ED discharge with a return visit for bronchiolitis within 12 hours, which results in hospital admission requiring supportive interventions for at least 12 hours. The choice of the first 12 hours is somewhat arbitrary, but is believed to have clinical face validity for representing the same episode of bronchiolitis and reflect premature discharge. A true need for hospital admission is also reflected by inpatient admission requiring the use of any of the major medical interventions for at least 12 hours for bronchiolitis patients in the hospital, including oxygen administration, intravenous fluid administration, suctioning for airway clearance, cardiovascular support, invasive positive pressure ventilation (mechanical ventilation), non-invasive positive pressure ventilation (NIPPV), chest physiotherapy, inhaled therapy (bronchodilator and mucolytics), and nutritional support (enteral feeding and total parenteral nutrition). Continuous positive airway pressure (CPAP), high-flow nasal cannula (HFNC) oxygen therapy, and bilevel positive airway pressure (BiPAP) are three examples of NIPPV.

3.1.2 Predicting hospital admission in the ED observation unit

About 36% of EDs in the U.S. have observation units treating patients for up to 24 hours [51]. The observation unit is used to avoid unnecessary or inefficient hospital admissions [52]. It offers treatments within defined areas for a period longer than the usual ED visit but shorter than a full inpatient hospital admission [52]. 55% of bronchiolitis patients in the observation unit fail discharge within 24 hours and are subsequently hospitalized [53]. Unexpected hospitalization from the observation unit involves transfer of care, and thus can decrease the efficiency and safety of both patient care and the observation unit itself [52].

At present, limited information exists on how to best select patients suitable for care in the observation unit [52]. For bronchiolitis patients in the observation unit, Yusuf et al. [54] identified several predictors of hospital admission: parental report of poor feeding or increased work of breathing, oxygen saturation < 93% at sea level, and ED treatment with racemic epinephrine and intravenous fluids. For all patients (not limited to bronchiolitis patients) in the observation unit, Alpern et al. [52] identified several predictors of hospital admission.
It remains an open problem to develop models to predict immediately before observation unit admission, whether a bronchiolitis patient will eventually have a true need for hospital admission. If one can predict that a bronchiolitis patient will have such a need, admission can be directed to the hospital rather than to the observation unit.

### 3.1.3 Predicting ICU admission

By applying a threshold to a single clinical parameter, Brooks et al. [55] built a model to predict whether an infant admitted to the general pediatric ward with RSV infection will be transferred to the ICU. In theory, this capability could support development of alternative care strategies, such as early use of NIPPV, for treating infants with RSV infection who present initially with milder symptoms. This study used previously healthy, full-term infants rather than all infants admitted with RSV infection. The developed model had a poor sensitivity below 30%, which is too low for identifying the majority of infants at risk. It remains an open problem to predict with high accuracy ICU transfer need for a patient admitted to the general inpatient ward with bronchiolitis (or RSV infection).

For children with bronchiolitis in the ED, Damore et al. [29] identified several predictors of ICU admission and built a logistic regression model to predict ICU admission, with an AUC of 80%. This study excluded bronchiolitis patients with previous ED visits. It used a small sample size of 50 ICU patients and hence could not assess less common risk factors.

For infants with RSV bronchiolitis, Mandelberg et al. [56] found peripheral blood mononuclear cell proliferation to be a risk factor for ICU admission. For children with bronchiolitis, García et al. [57-59] found three predictors of ICU admission: virus species (RSV vs. non-RSV), atelectasis/condensation, and co-infection. For children with RSV infection, Verger et al. [60-68] identified several risk factors for ICU admission: immature lung development, prematurity, chronic lung disease, congenital heart disease (defined as congestive heart failure, cyanosis, or pulmonary hypertension), neuromuscular impairment, high nasal RSV viral load, surfactant protein A2 polymorphism, age < 6 weeks, neurological disease, cerebral palsy, male gender, lung consolidation, lethargy, grunting, high arterial PaCO2, an ED visit in the past week, presence of moderate to severe retractions, inadequate oral intake upon presentation in the ED, and mental retardation.

Mansbach et al. [69] built a logistic regression model to predict whether a patient hospitalized with bronchiolitis will receive CPAP/intubation, with an AUC of 80%. CPAP/intubation was chosen as the outcome of interest, as it depends on better-defined objective criteria and thus is less variable than ICU admission.

Some ICU admissions are unnecessary. It is important to identify the true need for ICU admission rather than only the occurrence of ICU admission. Nevertheless, none of the work mentioned above has predicted the true need for ICU admission for bronchiolitis patients in the ED. This remains an open problem. A true need for ICU admission is defined as ED discharge.
with a return visit to the ED for bronchiolitis within 12 hours resulting in ICU admission, admission to the general ward with subsequent transfer to the ICU within 6 hours, or the use of ICU-level supportive interventions for at least 6 hours. Major medical interventions for bronchiolitis patients in the ICU include mechanical ventilation, NIPPV, cardiovascular support, and intensity of support beyond the capacity of the general inpatient ward (e.g., suctioning several times per hour).

3.1.4 Predicting optimal disposition in the primary care setting

63% of bronchiolitis patients have seen their primary care physicians during their illness prior to the ED visit [15]. The existing models for predicting hospital admission and ICU admission for bronchiolitis patients are mainly developed for the ED setting. For bronchiolitis patients who are younger than two years and seen in the primary care setting, Al-Shawwa et al. [70] identified three risk factors for hospitalization within ten days of evaluation: young age, passive smoking, and being RSV-positive. No other work has been published regarding predicting in the primary care setting, whether a bronchiolitis patient will use acute care (inpatient stay, urgent care, or ED visit) for bronchiolitis in the near future, such as within the next three days [20]. This is an open problem. By identifying the bronchiolitis patients at high risk for requiring acute care services in the near future, the primary care physicians can plan accordingly, such as arranging for an early follow-up primary care visit, making referrals to respiratory outpatient clinics [271], and providing more extensive counseling for caregivers. This could help bronchiolitis patients avoid future acute care usage, or seek it in a timely manner. A similar opportunity exists for making this prediction in the respiratory outpatient clinic setting as well as for a bronchiolitis patient on oxygen therapy at home.

3.1.5 Predicting safe discharge and unscheduled visit

Mansbach et al. [21] built a logistic regression model to predict whether a bronchiolitis patient in the ED will be safely discharged, with an AUC of 81%. A “safe discharge” was defined as a discharge to home without readmission in the following two weeks. This study excluded bronchiolitis patients with previous ED visits.

Norwood et al. [32] built a logistic regression model to predict whether a bronchiolitis patient discharged to home from the ED will have an unscheduled visit within two weeks of discharge. An “unscheduled visit” was defined as an urgent visit to an ED or clinic for worsening of bronchiolitis. By identifying children with bronchiolitis at high risk for unscheduled visits, the emergency physicians can devise a more personalized disposition plan, such as arranging for early follow-up primary care visits, and providing more extensive discharge counseling for parents and guardians [32]. This study excluded bronchiolitis patients with previous ED visits. The developed model had a low AUC of 64%.
Norwood et al. [32] reported that 17% of bronchiolitis patients discharged to home from the ED had unscheduled visits, 65% of which occurred within two days of the ED visit. It remains an open problem to develop a model to predict, upon ED discharge, whether a bronchiolitis patient will have an unscheduled visit in the very near future, such as in the next two days. If we could identify the bronchiolitis patients at high risk for an unscheduled visit in the very near future, a personalized care plan could be devised that might include admission to an observation or inpatient bed. Addressing predictions to the very near future ensures that the unscheduled visit will likely result from the same medical problem rather than from a new medical problem.

3.2 Making predictions related to respiratory support

3.2.1 Predicting whether a bronchiolitis patient will develop apnea

Apnea, a temporary suspension of breathing, is a life-threatening complication of bronchiolitis and occurs in 1.2% to 23.8% of children hospitalized with bronchiolitis [71, 72]. So far, little work has been done on predicting whether a bronchiolitis patient will develop apnea. The few published studies on this topic reported low prediction accuracy.

For infants hospitalized with RSV infection, Kneyber et al. [73] identified age below two months as a risk factor for apnea. For infants admitted to the ICU for RSV bronchiolitis, Schiller et al. [74] identified several risk factors for apnea: young age, low admission weight, low gestational age, admission from the ED, lack of hyperthermia, and no respiratory acidosis. For children hospitalized with bronchiolitis, Schroeder et al. [71] identified several predictors of apnea: corrected age of < 2 weeks, birth weight < 2.3 kg, history of apnea, preadmission respiratory rate of < 39 or > 70, and low room air oxygen saturation.

Willwerth et al. [75] developed three risk criteria for predicting whether an infant who is younger than six months and hospitalized with bronchiolitis will develop apnea. The risk criteria achieved a sensitivity of 100%, a low specificity of 64%, a low positive predictive value of 7%, and a negative predictive value of 100%.

It remains an open problem to predict with high accuracy whether a bronchiolitis patient in the ED, general inpatient ward, or ICU will develop apnea. By identifying bronchiolitis patients in the ED, general inpatient ward, or ICU who are at high risk for developing apnea, clinicians can make better decisions about monitoring, admission, and/or providing ventilatory support.

3.2.2 Predicting the use of a specific type of NIPPV

Depending on the hospital, the initiation and early titration of certain types of NIPPV, such as CPAP, may be restricted in the general inpatient ward [76]. Early initiation of NIPPV on bronchiolitis patients needing it is important for improving patient outcomes [77]. If one can predict which bronchiolitis patients in the ED or general inpatient ward will need a specific type of
NIPPV, admission could be directed to an inpatient unit that can deliver this type of NIPPV [76]. This will help reduce their need for in-hospital transfers, mechanical ventilation, and ICU admissions.

So far, Evans et al. [76] is the only study along this direction. For bronchiolitis patients in the ED, that study identified several predictors of using CPAP: oxygen requirement in the ED, low oxygen saturation, young age, high respiratory rate, high heart rate, low Glasgow Coma Scale score, and low gestational age.

It remains an open problem to develop models to predict whether a bronchiolitis patient in the ED or general inpatient ward will eventually use a specific type of NIPPV, such as CPAP, HFNC, or BiPAP. In constructing such predictive models, contraindications to NIPPV support, such as those listed in Mayordomo-Colunga et al. [78], should be considered.

3.2.3 Predicting NIPPV failure

Bronchiolitis represents the largest cohort of children treated with NIPPV [79]. NIPPV has several advantages over mechanical ventilation [79, 80]. First, by avoiding intubation, NIPPV reduces patient risk of airway damage, ventilator-associated pneumonia, and other nosocomial infections. Second, NIPPV can help decrease patients’ requirement for sedation, along with its costs and risks. Third, NIPPV increases patient mobility and allows some patients to be managed outside of an ICU. Fourth, NIPPV allows greater flexibility in approach to nutritional support. Thus, mechanical ventilation should be avoided if a patient can be successfully managed with NIPPV. For bronchiolitis patients in the ICU, selectively replacing mechanical ventilation with NIPPV has been shown to reduce the rate of ventilator-associated pneumonia, duration of supplemental oxygen, and hospital length of stay [81].

NIPPV works for only a subset of patients requiring respiratory support. As reported in the literature, 5-43% of patients failed NIPPV and subsequently required mechanical ventilation [78, 80-83]. So far, no guideline for NIPPV has been published for children [82]. It is often not intuitively obvious which pediatric patients will benefit from NIPPV [79].

Various predictors of NIPPV failure have been proposed in the research literature [79]. For infant patients with bronchiolitis in the ICU, Abboud et al. [80] identified pre-HFNC PCO₂ and respiratory rate as predictors of HFNC failure. For pediatric patients younger than two years who received HFNC within 24 hours of initial triage in the ED, Kelly et al. [83] identified three predictors of HFNC failure: triage respiratory rate > 90th percentile for age, initial venous PCO₂ > 50 mm Hg, and initial venous pH < 7.3. 46% of these patients had bronchiolitis. For pediatric patients in the ICU, Mayordomo-Colunga et al. [78, 82] identified several predictors of NIPPV failure: a fraction of inspired oxygen (FiO₂) > 80% after 1 hour of NIPPV, type 1 acute respiratory failure, high Pediatric Risk of Mortality score, and low respiratory rate decrease (at 1 hour and at 6 hours).
It remains an open problem to develop models to predict whether a bronchiolitis patient in the hospital requiring positive pressure ventilation support will fail NIPPV. If one can predict which bronchiolitis patients in the hospital will fail NIPPV, patients can be put on mechanical ventilation immediately rather than be subject to the stress involved in a failed attempt at NIPPV [84]. Inappropriate delay of mechanical ventilation can cause clinical deterioration and increase morbidity and mortality [79]. This prediction would ideally be made before a patient is put on NIPPV. Also, after a patient is put on NIPPV, we should continuously monitor the patient and predict NIPPV failure from time to time [79]. This can help minimize inordinate delay in mechanical ventilation when it is the best therapeutic approach.

The failure rate of NIPPV varies for various diseases [79]. Also, each type of NIPPV, such as CPAP, HFNC, or BiPAP, has its own properties. To maximize the prediction accuracy of NIPPV failure in bronchiolitis patients, we should develop a predictive model specifically for bronchiolitis rather than use a generic predictive model for all diseases and all types of patients. Moreover, for each type of NIPPV, there may be a need to develop a predictive model specifically tailored to it [85].

3.2.4 Predicting the use of supplemental oxygen

For patients hospitalized with bronchiolitis, García et al. [57, 86, 87] identified several predictors of using supplemental oxygen: household tobacco smoking, cyanosis, sternal retraction, intercostal recession, chronic lung disease, trisomy 21, congenital heart disease, virus species, and prematurity. For bronchiolitis patients in the ED and inpatient ward, McCallum et al. [88] used a bronchiolitis severity score to predict the use of supplemental oxygen at 12 and 24 hours, with a low AUC of 68% and 75%, respectively.

It remains an open problem to develop models to predict with high accuracy whether a bronchiolitis patient in the ED or inpatient ward will eventually use supplemental oxygen. A related open problem is to predict the number of days for which a bronchiolitis patient in the ED or inpatient ward will use supplemental oxygen. This will help develop care processes stratified by the predicted need of supplemental oxygen. For example, if a bronchiolitis patient in the ED is predicted to use supplemental oxygen for no more than four days, the patient may be discharged from the ED.

3.2.5 Predicting the use of mechanical ventilation

Mechanical ventilation is used on 7-21% of infants hospitalized with RSV bronchiolitis [89]. For patients hospitalized with bronchiolitis, Garcia et al. [57, 86, 90, 91] identified several predictors of using mechanical ventilation: weight, prematurity, household tobacco smoking, young age, female gender, virus species, failure to thrive, underlying disease, and pneumonic infiltration on chest X-ray. For previously healthy patients hospitalized with RSV infection, DeVincenzo et al. [92] found
weight and RSV viral load to be predictors of using mechanical ventilation and ICU admission. For patients with RSV bronchiolitis, Brand et al. [93] found interleukin-8 level, chemokine (C-C motif) ligand 5 level, and CD4+ T-cell count to be predictors of using mechanical ventilation. For patients hospitalized with RSV infection, Verger et al. [60-64, 94-97] identified several predictors of using mechanical ventilation: interferon-γ level, cardiac troponin I obtained in the ED, congenital hydrocephalus without spina bifida, choanal atresia, lung agenesis, hypoplasia or dysgenesis, cleft lip/palate, immature lung development, chronic lung disease, congenital heart disease, neuromuscular impairment, surfactant protein A2 polymorphism, age < 6 weeks, neurological disease, and Down syndrome.

For previously healthy patients with RSV infection, Prodhan et al. [40] used several variables collected in the ED to build a logistic regression model to predict the use of mechanical ventilation in the ICU. The abnormality of certain variables is based on age-specific norms. On the training set, the model achieved an AUC of 91.5%, a sensitivity of 71%, a specificity of 96%, a positive predictive value of 86%, and a negative predictive value of 91%. However, the model was not evaluated on a validation set. Usually, a predictive model’s performance measures obtained from a validation set are worse than those obtained from the training set. It remains an open problem to develop a highly reliable model to predict the need of mechanical ventilation in bronchiolitis patients.

### 3.2.6 Predicting prolonged mechanical ventilation

For infants who are admitted to the ICU with RSV infection and require mechanical ventilation, Prodhan et al. [98, 99] listed several risk factors for prolonged mechanical ventilation (>8 days). Using various variables including some collected during the first two days after intubation, Prodhan et al. [98] built a logistic regression model to predict prolonged mechanical ventilation, with a large AUC of 92% and a reasonably good accuracy of 84%.

It remains an open problem to develop models to predict prolonged mechanical ventilation on a patient who is admitted to the ICU with bronchiolitis and requires mechanical ventilation.

### 3.2.7 Predicting the use of extracorporeal membrane oxygenation

Among ventilated children with RSV bronchiolitis in the ICU, 9.3% of them will fail mechanical ventilation, develop profound hypoxemia, and need extracorporeal membrane oxygenation support [99]. For ventilated infants with RSV bronchiolitis in the ICU, Flamant et al. [99] identified chronic lung disease as a predictor of using extracorporeal membrane oxygenation. It remains an open problem to develop models to predict whether a patient hospitalized with RSV bronchiolitis...
will need extracorporeal membrane oxygenation support. If one can predict which patients with RSV bronchiolitis will need extracorporeal membrane oxygenation support, patients can be given this support expeditiously to help improve their outcomes.

3.2.8 Predicting extubation failure

In the ICU, about 40% of patients need mechanical ventilation for sustaining their lives [100]. 52% of complications are related to ventilator use [101], particularly if it is prolonged [102]. Prolonged ventilator use can cause subglottic injury, respiratory infections, and chronic lung disease [100]. To reduce cost and risk of ventilator-induced lung injury, nosocomial pneumonia, airway trauma from the endotracheal tube, and unnecessary sedation, it is important to extubate patients as soon as possible [100]. However, this needs to be done carefully. Premature extubation can cause respiratory muscle fatigue, gas exchange failure, and loss of airway protection [100]. It can also increase morbidity, mortality, ICU stay, mechanical ventilation duration, as well as the risk of nosocomial infection, sedative dependency, and airway trauma due to re-intubation [103-106].

Premature extubation usually leads to extubation failure, which refers to the need to re-intubate and restore mechanical ventilation within a certain time period, such as 48 hours, after extubation [104-108]. Extubation failure happens to 15-20% of child patients [104, 106-108], 22-28% of premature neonatal patients [104], and 17-19% of adult patients [104] in the ICU.

Due to their different levels of experience and preferences, clinicians have variable performance in determining an appropriate extubation time [109]. To help clinicians better identify the earliest time that a specific patient can be safely extubated, it is desirable to develop models to predict extubation failure [100]. Such predictive models should use multiple indices obtained from diverse instruments and modalities, because an index obtained from a single device can be easily affected by systematic error that cannot be reduced through increasing the sample size [100]. Although various predictors of extubation failure have been proposed in the research literature [100, 104-109], there is currently no consensus on which predictors should be used [100, 107].

Using support vector machine, Hsu et al. [100] developed a machine learning model to predict extubation failure in adult patients, with an accuracy of 91%. A clinical decision support tool based on the predictive model was adopted for adult patients on mechanical ventilators in a respiratory care center. On average for each such patient, the adoption shortened mechanical ventilation duration by 5.2 days and reduced healthcare cost by US$1,500 [110].

Mueller et al. [109] developed several machine learning models to predict extubation failure in neonatal patients. The best of these models achieved an AUC of 76%, which is insufficient for routine clinical use [109].

The predictive models developed in Hsu et al. [100] and Mueller et al. [109] are unlikely to work well for bronchiolitis patients. Certain predictors of extubation failure in adults are useless in infants and children [104]. Also, the predictors of
extubation failure can vary for patients with different diseases [107]. To maximize the prediction accuracy of extubation failure in bronchiolitis patients, we should develop a predictive model specifically for bronchiolitis rather than use a generic predictive model for all diseases and all types of patients.

For a single ventilation mode, the pressure control mode, Johnston et al. [107] identified several predictors of extubation failure in bronchiolitis patients in the ICU: weight ≤ 4 kg, tidal volume ≤ 4 mL/kg, minute volume ≤ 0.8 mL/kg/min, maximal inspiratory pressure ≤ 50 cm H$_2$O, load/force balance ≥ 5, and rapid shallow breathing index ≥ 6.7. This was done using a small data set including 40 infant patients, six of whom experienced extubation failure. The predictor of maximal inspiratory pressure ≤ 50 cm H$_2$O achieved an AUC of 97%. It remains an open problem to develop models to predict extubation failure in bronchiolitis patients in the ICU for all ventilation modes. To obtain reliable performance measures on them, such predictive models should be evaluated on large data sets.

3.3 Making predictions about the value of tests and other evaluations on a bronchiolitis patient

3.3.1 Predicting serious bacterial infection

Delayed diagnosis and treatment of a serious bacterial infection can have a large impact on long-term health [111]. To avoid such a delay, sepsis/meningitis evaluation is frequently performed on patients hospitalized with bronchiolitis. As mentioned in Antonow et al. [112], 49.6% of infants who are younger than 60 days and hospitalized with bronchiolitis receive sepsis evaluation. However, most of these evaluations are unnecessary. In fact, only 2.2% of febrile infants hospitalized with bronchiolitis have serious bacterial infections [113]. For children hospitalized with RSV infection, this rate is 1.6% [114].

Sepsis/meningitis evaluation incurs significant cost, discomfort to a child, and stress to the child’s family [114]. As reported in Antonow et al. [112], among infants hospitalized with bronchiolitis, those who underwent sepsis evaluation had a hospitalization cost of $4,507 and a hospital length of stay of 3.4 days on average. In contrast, the others had a lower hospitalization cost of $2,998 and a shorter hospital length of stay of 2.8 days on average. Thus, it is desirable to eliminate unnecessary sepsis/meningitis evaluations.

For children hospitalized with RSV infection, Bloomfield et al. [115] identified three risk factors for bacteremia: nosocomial RSV infection, cyanotic congenital heart disease, and ICU admission. It remains an open problem to develop models to predict serious bacterial infection for patients hospitalized with bronchiolitis. If we predict that a bronchiolitis patient is unlikely to have a serious bacterial infection, we can eliminate unnecessary sepsis/meningitis evaluation on the patient.

3.3.2 Predicting an unhelpful chest X-ray
Upon admission, a chest X-ray is performed on 72% of children hospitalized with RSV infection [11]. 37% of those patients have a normal chest X-ray [116]. The chest X-ray does not influence clinical care for 92.8% of those patients [117]. If one can identify the patients who will have an unhelpful chest X-ray beforehand, their exposure to radiation and costs associated with unnecessary chest X-rays can be minimized. Since patients having a chest X-ray are more likely to receive inappropriate additional therapies such as antibiotics, saving an unhelpful chest X-ray can also save these unnecessary therapies, if any, and their associated costs [28, 117].

To help clinicians decide which children with RSV infection require no chest X-ray, Kneyber et al. [116] developed a logistic regression model to predict a normal chest X-ray in children with RSV infection, with an AUC of 80%. About 25% of patients with an abnormal chest X-ray were falsely predicted as having a normal chest X-ray.

It remains an open problem to develop models to predict an unhelpful chest X-ray for bronchiolitis patients. To reduce the manual labeling overhead of obtaining a large enough training set needed for training such predictive models, natural language processing or information extraction [118] can be conducted on historical chest X-ray reports [119] of bronchiolitis patients to facilitate labeling previously-performed chest X-rays as helpful or unhelpful.

### 3.3.3 Identifying RSV infection

Early identification of RSV infection is important for several reasons [120]. First, RSV is a highly contagious organism [60]. During epidemics, up to 50% of inpatients can be affected by RSV [121]. The longer a patient stays in a hospital, the more likely he/she will be affected by RSV [121]. Hence, among hospitalized patients, we should identify those with RSV infection and group them properly to minimize nosocomial infection [122]. Second, well-appearing infants with RSV infection have a low risk for serious bacterial infection [123-125]. If we know that an infant has RSV infection, but is otherwise well-appearing, we may eliminate unnecessary sepsis/meningitis evaluation on the infant [113]. Third, infants with RSV infection require careful consideration for prognostication and disposition purposes.

As mentioned in Durani et al. [120], the definitive diagnosis of RSV infection is based on tissue culture and often delayed. Rapid antigen test for RSV has low sensitivity and may be unavailable in developing countries. Polymerase chain reaction (PCR) testing for RSV is both sensitive and specific. However, its turnaround time may be excessive. Also, it may be unavailable in many settings including developing countries. Thus, it is desirable to develop classification models to help clinicians identify RSV infection cases. Such a classification model can be used by itself to reduce unnecessary rapid antigen and PCR tests for RSV. It can also be combined with rapid antigen or PCR tests for RSV to increase the tests’ predictive value.
For child patients in a suburban hospital suspected by an ED physician of having RSV infection, Durani et al. [120] built a logistic regression model to identify RSV infection. This model achieved a low AUC of 66%, a sensitivity of 80%, a low specificity of 68%, a positive predictive value of 82%, and a low negative predictive value of 66%. For infants hospitalized for acute lower respiratory tract disease, Riccetto et al. [126] identified several risk factors for RSV infection: gestational age < 35 weeks, birth weight $\leq$ 2.5 kg, mother has less than five years of school education, and pulse oximetry < 90% upon hospital admission time. It remains an open problem to develop classification models that can accurately identify RSV infection and provide reliable prediction for all patient populations suspected of having RSV infection. Similar models could be developed for other infectious agents.

3.4 Making predictions regarding the current hospitalization

3.4.1 Predicting hospital length of stay

Researchers have developed several models to predict hospital length of stay for bronchiolitis patients. By identifying the bronchiolitis patients at high risk for a prolonged hospital stay, we can provide them with early interventions [37] to reduce this risk. Except for the model developed in Weisgerber et al. [37], all other models make predictions upon hospital admission time. We describe these predictive models one by one as follows.

Corneli et al. [43] used CART to predict whether an infant with bronchiolitis will have a hospital stay longer than one night, with a sensitivity of 77% and a low specificity of 57%. This study used various criteria to exclude most infants with bronchiolitis in the ED.

Weisgerber et al. [37] used CART to predict whether a bronchiolitis patient will have a prolonged hospital stay. Prediction was made at two days after hospital admission, with a low AUC of 72%. 70% of infants admitted to the Children’s Hospital of Wisconsin with bronchiolitis were placed on the hospital’s bronchiolitis treatment protocol. Prediction was made only for the bronchiolitis patients who were placed on the protocol and had no comorbidity that might impact hospital length of stay.

Moler and Ohmit [127] built a logistic regression model to predict whether a patient with RSV infection will have a prolonged hospital stay. Kneyber et al. [128] showed that this predictive model had a low AUC of 65%, and failed to identify a considerable number of patients with a prolonged hospital stay.

Walsh et al. [24] built a logistic regression model to predict whether a bronchiolitis patient will have a hospital stay longer than the mean hospital length of stay, but did not show the model’s prediction accuracy. Walsh et al. [36] used neural network ensemble to predict the actual hospital length of stay, with a low accuracy of 65%.
For patients hospitalized with RSV infection, Verger et al. [60-64, 67, 92, 96, 129-132] identified several predictors of a prolonged hospital stay: congenital anomaly (congenital heart disease), low weight, high nasal RSV load, male gender, underlying cardiac or respiratory disease or anomaly, the need for mechanical ventilation, congenital hydrocephalus without spina bifida, choanal atresia, lung agenesis, hypoplasia or dysgenesis, cleft lip/palate, immature lung development, chronic lung disease, prematurity, immunodeficiency, malformations of the esophagus, Down syndrome, surfactant protein A2 polymorphism, parental smoking, age < 6 weeks, neurological disease, and ICU admission. For infants hospitalized with RSV bronchiolitis, Kott et al. [133, 134] identified several predictors of hospital length of stay: age, urinary cysteinyl leukotriene E4 concentration, and impaired plasma tumor necrosis factor α and interleukin-6 production capacity. For patients hospitalized with bronchiolitis, Hervás et al. [58] found virus species (RSV vs. non-RSV) to be a predictor of hospital length of stay.

It remains an open problem to predict with high accuracy hospital length of stay of a bronchiolitis patient.

3.4.2 Predicting hospitalization cost

For an individual patient to be admitted for RSV infection, Rietveld et al. [135] developed a linear regression model to predict hospitalization cost, with an $R^2$ of 8% indicating low prediction accuracy. Fieldston et al. [136] showed that zip code-based median annual household income is a predictor of hospitalization cost for bronchiolitis patients. It remains an open problem to predict with high accuracy hospitalization cost of a patient to be admitted for bronchiolitis.

3.5 Making predictions related to the use of palivizumab

3.5.1 Predicting RSV hospitalization

RSV infection consumes a lot of healthcare resources. Each year, RSV infection causes more than 90,000 pediatric hospitalizations and 4,500 deaths in the U.S. [137], and 160,000 deaths globally [138]. In the U.S., the cost of one RSV hospitalization varies between $3,777 and $13,241 [139]. The total cost of all RSV hospitalizations was $1.1 billion in 2002 [139]. Among pediatric patients admitted through the ED, more than 74% of RSV infection cases are diagnosed as bronchiolitis [138].

At present, palivizumab is the only product approved for preventing RSV lower respiratory tract infection [140, 141]. Palivizumab is shown in studies to be safe and effective. For example, its use can reduce RSV hospitalization by 78% in children born at 35 weeks gestation or less [137].

Palivizumab is expensive, with a cost of a course of treatment exceeding the cost of a single uncomplicated RSV admission. On average, immunizing one patient with palivizumab over the winter season costs $3,688-6,140 in the U.S., £2,544-4,235 in
the U.K., €3,969-6,607 in various European countries [142], and CAN$6,540 [143] in Canada. To be financially viable, palivizumab is provided to only those children at high risk for complicated RSV hospitalization rather than to all children [144]. To help clinicians identify the best candidates for palivizumab immunization, a model is often used to predict the likelihood that a specific child will experience RSV hospitalization in the next few months [144].

Most of the existing predictive models for RSV hospitalization were developed for preterm infants born at 33-35 weeks gestation. Each year, 3-5% of annual births belong to this cohort [145]. If no preventative treatment is used, 2-10% of infants in this cohort will experience RSV hospitalization [146], which is associated with a large increase in subsequent healthcare resource usage and mortality [147].

For infants in this cohort, Simões et al. [144] used discriminatory function analysis to build a model to predict RSV hospitalization. On various data sets, the predictive model achieved an AUC of 63-79%, a classification accuracy of 66-77%, a sensitivity of 46-75%, a specificity of 67-79%, a positive predictive value of 10-75%, and a negative predictive value of 73-96% [138, 144, 148-150].

Similarly, Sampalis et al. [146] built a logistic regression model to predict RSV hospitalization. On various data sets, the predictive model achieved an AUC of 70-76%, a sensitivity of 61-68%, a specificity of 66-72%, a false positive rate of 34%, and a false negative rate of 39%.

Over 85% of all children experiencing RSV hospitalization are born at term or later [33, 135]. As shown in Paes et al. [151], the model developed in Sampalis et al. [146] cannot accurately predict RSV hospitalization or ED visit for term infants. Also, the predictors of RSV hospitalization that work for preterm infants born at 33-35 weeks gestation have little or no predictive power for term infants.

Palivizumab is administered on a monthly basis [137]. Rietveld et al. [33] built a logistic regression model to estimate a young child’s monthly risk of RSV hospitalization, with an AUC of 80%. If one can accurately estimate this monthly risk, a child at high risk for RSV hospitalization can be given palivizumab during only selected months when his/her risk is higher, rather than all months, in the winter. This will reduce immunization cost.

It remains an open problem to develop models to predict the true need for (complicated) seasonal RSV hospitalization for all young children, not only those born at 33-35 weeks gestation [151]. To build such predictive models, we can use the predictors of RSV/bronchiolitis hospitalization that have already been identified in the literature for young children born at various gestational stages [60-63, 66, 96, 131, 141, 144, 145, 152-159, 161-182]. We can consider certain chronic conditions, such as congenital heart disease, chronic lung disease, and severe neurological diseases, which may increase the risk of RSV
hospitalization, but are ignored in existing predictive models [144]. We can make prediction on a monthly basis, in a way similar to that in Rietveld et al. [33], accounting for community viral prevalence, prevailing weather patterns [183-185], etc. Since each gestational stage has its own set of predictors of RSV hospitalization [151], we should build a separate predictive model for each gestational stage tailored to children born at that stage. Moreover, RSV hospitalization cost varies by several times in infants with different risk factors and different pre-existing medical conditions [139]. We can predict a young child’s expected RSV hospitalization cost based on the probability of RSV hospitalization estimated for the child. Then we could use the predicted cost and the potential benefit to optimize the use of palivizumab.

Another open problem, which is related to the one mentioned above, is to develop models to predict the true need for (complicated) bronchiolitis hospitalization for young children. By identifying children at high risk for experiencing bronchiolitis hospitalization, physicians can target preventive and monitoring strategies, such as immunization with influenza vaccine, towards them.

3.5.2 Predicting compliance with palivizumab

To maintain a serum concentration level of palivizumab that is sufficient for protecting against RSV, a child at high risk for RSV hospitalization must receive multiple doses of palivizumab throughout the RSV season, usually one dose of intramuscular injection per month [186]. If a high-risk child misses one or more doses of prescribed palivizumab, he/she could experience hospitalization with high cost, and the value of the high expense already incurred for previously administered doses would be diminished [187]. Thus, it is important to ensure patient compliance with palivizumab.

The rate of noncompliance with palivizumab varies across different care settings: 70% for physician offices, 26% for day health centers, 12% for pulmonologists’ offices, 11% for outpatient clinics, and 9% for at-home administration of palivizumab by a visiting nurse [187]. Known risk factors for noncompliance with palivizumab include parental smoking, Medicaid enrollment, minority race, low socioeconomic level, low parental expectations for the benefits of RSV prophylaxis, lack of transportation, and language difficulties [186, 187]. At present, multiple models exist for predicting medication compliance in patients with various diseases such as heart failure [188]. However, none of these models was built specifically for predicting compliance with palivizumab.

It is desirable to develop models to predict compliance with palivizumab. This is an open problem. If a child is predicted to be unlikely to be compliant with palivizumab, we can adopt one or more interventions to increase his/her likelihood of being compliant. Examples of such interventions include reminder phone calls, comprehensive multidisciplinary programs including extensive counseling of parents, calendars with sticker reminders, multilingual call center, and administering palivizumab in a
care setting that is known to have a high compliance rate [187]. When choosing interventions, we should consider both the intervention’s cost and the child’s predicted likelihood of being noncompliant. Expensive interventions, such as at-home administration of palivizumab by a visiting nurse, should be reserved for those patients with the highest predicted likelihood of being noncompliant.

3.5.3 Predicting community RSV activity

RSV infections primarily occur in the late fall, winter, and early spring in the U.S. and other countries with temperate climates [189]. However, community RSV activity varies substantially by year and by location [189]. Even during the same year, the onset of community RSV activity can vary between communities in close proximity [190].

The onset of RSV season is defined as the time when positive antigen tests are greater than 10% of those sampled [60]. Predicting the onset and duration of a local community’s RSV activity can help improve the cost-effectiveness of palivizumab [191]. Ideally, to ensure that at-risk children are adequately protected from RSV infection, the first dose of palivizumab should be administered before the onset of community RSV activity [190, 192]. To avoid the cost of unnecessarily administering palivizumab, the last dose of palivizumab should be administered before the offset of community RSV activity [190].

It has been shown that community RSV activity, such as the weekly number of RSV hospitalizations, is significantly associated with the weekly mean and minimum temperature, water vapor pressure, relative humidity, barometric pressure [191], ultraviolet B radiance [183], and the weekly mean PM$_{10}$ concentration [193], often with a time lag. Thus, environmental factors can be used to predict community RSV activity.

Along this direction, Walton et al. [184] built a naive Bayes model to predict the onset of community RSV activity up to three weeks in advance, with a sensitivity of 67% and a specificity of 94%. The model’s prediction accuracy is inadequate for practical use [185]. du Prel et al. [194] developed an autoregressive integrated moving average model to predict the biweekly number of RSV hospitalizations, with an $R^2$ of 65% indicating low prediction accuracy.

Zachariah et al. [195] identified three predictors of the duration of the RSV season: the number of children younger than five years per room, the number of children per square kilometer, and whether the location is urban or rural. It remains an open problem to predict community RSV activity with high accuracy.

3.6 Predicting whether a bronchiolitis patient will later be diagnosed with asthma

14% to 40% of bronchiolitis patients will eventually have asthma [25, 196], with the association persisting into adulthood [196-203]. In young children, asthma is heralded by acute bronchiolitis in 90% of cases, one third of which require
hospitalization [204]. It is desirable to develop a model to predict whether a bronchiolitis patient will later be diagnosed with asthma [205]. This will likely involve modeling diverse coincident variables unrelated to bronchiolitis. By identifying the bronchiolitis patients at high risk for later being diagnosed with asthma and by scheduling more frequent follow-up visits for them, we can obtain several benefits. First, we can detect asthma and subsequently start asthma treatment on these patients earlier [206]. This will help them avoid future asthma exacerbations, such as hospitalization and ED visits, which are undesirable and often expensive. Second, we can consider risk modulators such as using palivizumab and emerging novel immunomodulatory therapies on these patients to help prevent RSV infections. Third, we can route these patients into future randomized clinical trials to advance bronchiolitis and asthma prevention research [207]. Fourth, by distinguishing early asthma patients from transient early wheezers, clinicians can better manage early wheezers [208]. This helps avoid under-treatment with a risk of complications, such as permanent impairment of lung function, as well as over-treatment with potentially harmful medicines, such as steroids in transient early wheezers [196].

For bronchiolitis patients, various predictors of recurrent wheezing and emerging asthma have been identified in the research literature [171, 180, 197, 202, 204, 206-233]. Using clinical variables assessed at age two, Mikalsen et al. [205] built a logistic regression model to predict asthma diagnosis at age 11. The model achieved a low sensitivity of 65%, a specificity of 82%, a low positive post-test probability around 50%, and a low negative post-test probability around 11%. At present, no predictive model for asthma diagnosis can attain satisfactory accuracy [206, 207].

It remains an open problem to develop models to predict with high accuracy whether a bronchiolitis patient will later be diagnosed with asthma. To build such predictive models, we should use the risk factors for asthma known in the existing literature [234-239] rather than only those for bronchiolitis patients. As one predictive model does not fit all, we should develop separate predictive models for children presenting with bronchiolitis at <6, 6-12, and 13-24 months of age [240]. Since characteristics associated with future asthma diagnosis after bronchiolitis vary by age [225], we should choose a targeted age and build predictive models specifically for it.

### 3.7 Making other predictions

In addition to the predictions mentioned above, researchers have worked on making other predictions related to bronchiolitis. Miller et al. [160, 241-249] listed several risk factors for severe bronchiolitis or severe RSV infection, such as age < 6 weeks at presentation, apnea, preterm birth, lung disease (e.g., chronic lung disease, cystic fibrosis), congenital heart disease, congenital or acquired immunodeficiency, multiple congenital abnormalities, and severe neurological disease. Shaw et al. [34] found several predictors for identifying infants with bronchiolitis at risk for having more severe disease: ill or toxic general
appearance, oxygen saturation < 95% at sea level, gestational age < 34 weeks, respiratory rate > 70/min, atelectasis on chest roentgenograms, and postnatal age < 3 months. For previously healthy term infants admitted to the ED and ICU for bronchiolitis, Papoff et al. [250] identified three predictors of requiring mechanical ventilation or NIPPV: age < 30 days, RSV infection, and blood lymphocyte counts < 3200/µL.

Flaherman et al. [251] identified several predictors of bronchiolitis episode of care in infants ≥32 weeks gestation: male gender, African-American and Hispanic race/ethnicity, low gestational age, having at least one sibling < 5 years of age, congenital anomaly, family history of asthma, degree of oxygen exposure during the birth hospitalization, and chronic lung disease. For patients hospitalized with RSV lower respiratory tract infection, Wang et al. [252] identified several predictors associated with ICU admission, mechanical ventilation, and hospital length of stay. A history of apnea, hypoxemia on admission, and pulmonary consolidation were associated with all three outcomes. Aboriginal race, age < 6 weeks, and underlying pulmonary disease were associated with the first two outcomes. Prematurity, immunosuppression, and congenital heart disease were associated with hospital length of stay.

For patients hospitalized with RSV infection, Akiyama et al. [253] identified both obesity and leanness as risk factors for a prolonged disease duration. For patients hospitalized with RSV bronchiolitis, Howidi et al. [242] found age to be a predictor of the duration of oxygen therapy. For infants with RSV infection in the ICU requiring mechanical ventilation, Tasker et al. [254] found mean airway pressure and alveolar-arterial oxygen gradient measured in the first two days of mechanical ventilation to be predictors of ICU length of stay.

Spaeder and Fackler [255] built a time series model to predict the number of viral respiratory illness cases that will present to a pediatric ICU. Bronchiolitis in infancy is a prototypic example of such a viral respiratory illness.

Houben et al. [30] built a logistic regression model to predict whether a healthy term newborn will develop RSV bronchiolitis, with an AUC of 72%. By identifying children at high risk for developing RSV bronchiolitis, physicians can target preventive and monitoring strategies towards them. Houben et al. [256] identified low amniotic fluid interleukin-8 and low tumor necrosis factor-α as risk factors for RSV bronchiolitis in healthy term infants.

3.8 Personalized recommendation of treatment models

Many supportive treatments and drug interventions are routinely applied to bronchiolitis patients. Examples of such supportive treatments include bronchodilator use, supplemental oxygen, intravenous hydration, parenteral antibiotic, systemic corticosteroid, intensive care, mechanical ventilation, and respiratory physiotherapy [11]. Despite many years of research, physicians are still unclear about how to optimally manage bronchiolitis patients in various care settings (ED, hospital, and
outpatient), including which treatment model is the optimal one for a specific bronchiolitis patient [2, 4, 10, 241]. As a result, many bronchiolitis management decisions are made subjectively and possibly sub-optimally.

Physicians apply treatment models to manage patients. Treatment models often address the type, sequence, and/or intensity of the treatments used. Differing bronchiolitis patients may have different optimal treatment models [241]. In particular, a treatment model may have no clear benefit for all children with bronchiolitis, but could be useful for a subset of children with bronchiolitis [241].

To support personalized care, we recently proposed using machine learning techniques and treatment model profile predictors to identify an effective treatment model for a patient tailored to the patient’s characteristics [257]. A treatment model’s profile includes historical data of all patients managed by the treatment model. An example treatment model profile predictor is the logarithm of the number of patients who are of a particular type (e.g., gender) and have been managed by the treatment model. In our proposed method, for each treatment model, we predict the patient’s outcome if the patient is going to be managed by the treatment model. This requires answering intervention queries through performing causal inference [258, 259]. For instance, we can use a matching method to deal with the effects of confounders and conduct case-based reasoning [258, 260]. Then all treatment models are sorted according to the predicted patient outcome. The higher-ranked treatment models are more likely to help the patient achieve a good outcome.

As is the general case with machine learning, the prediction results are not 100% accurate. Consequently, the top-ranked treatment models may or may not be the optimal ones for the patient. Nevertheless, if the prediction results are reasonably accurate, we would expect that the suitability of the top-ranked treatment models for the patient will be close to that of the optimal treatment model for the patient, or at least this method can help us avoid extremely bad treatment models for the patient with a high probability [261].

Our previous work [257] proposed only a high-level idea for matching patients with treatment models in general without going into any detail. To make this idea work for bronchiolitis, much work remains to be done, such as designing predictors and a matching method tailored to bronchiolitis. We would expect that after proper extension, machine learning and causal inference techniques that will be developed for personalized recommendation of treatment models for bronchiolitis can also be used for some other healthcare applications, such as personalized search for individual healthcare providers [257, 261].

**Opportunities for improving predictive modeling for bronchiolitis**

In the above, we have mentioned multiple opportunities for developing predictive models for various issues related to bronchiolitis for which no predictive model has ever been built. There are several other opportunities for improving predictive
modeling for bronchiolitis, in the categories of making predictions for all bronchiolitis patients and improving prediction accuracy.

3.9 Making predictions for all bronchiolitis patients

Almost every existing predictive model for bronchiolitis was developed using some non-empty exclusion criterion for the patient population, such as excluding bronchiolitis patients with previous ED visits. It would be desirable to develop additional predictive models to cover all bronchiolitis patients. All other things being equal, predictive models that apply to more patients are more useful.

3.10 Improving prediction accuracy

Most of the existing predictive models for bronchiolitis have inadequate accuracy, exemplified by an AUC (much) less than 80%. It remains an open problem to improve the accuracy of the predictions made for bronchiolitis. There are several potential ways to do this, including using lab test data, using information embedded in clinical text, using machine learning methods, using physician practice predictors, using treatment model profile predictors, and using large data sets and exhaustive variable sets. We describe these ways one by one as follows.

3.10.1 Using lab test data

With rare exceptions such as Marlais et al. [31, 40, 42], existing predictive models for bronchiolitis use no lab test data. As is the case with clinical predictive modeling in general, we would expect that using lab test data can help improve prediction accuracy. For predictive modeling for bronchiolitis, potentially useful lab test data include viral test results (e.g., viral co-infection), complete blood count, capillary blood gas, blood lactate level, chest X-ray result, and basic metabolic profile result. While no benefit has yet been recognized from pursuing lab and X-ray evaluation in bronchiolitis patients, in a new modeling approach, it could be beneficial to add these lab test data as predictors into the predictive models for bronchiolitis.

3.10.2 Using information embedded in clinical text

Much clinical information is unavailable from structured data, but resides in unstructured clinical text. However, almost none of the existing predictive models for bronchiolitis use information embedded in clinical text. To improve the predictive models’ accuracy, we can use medical information extraction techniques to automatically extract information from clinical text [118], such as by mapping chief complaint strings to syndrome categories [262]. The extracted information is added as predictors into the predictive models for bronchiolitis.
3.10.3 Using machine learning methods

With rare exceptions such as Walsh et al. [36, 37, 43], almost all existing predictive models for bronchiolitis are based on the statistical method of logistic regression. As is the case with predictive modeling in general, machine learning methods such as support vector machines and random forests often achieve higher prediction accuracy than logistic regression [38]. It would be interesting to compare various machine learning methods for predictive modeling for bronchiolitis.

Much existing work on predictive modeling for bronchiolitis, such as Parker et al. [2, 21, 29-31, 42], starts from univariate logistic regression to identify useful predictors. This approach may miss an opportunity to identify combinations of predictors that together have good predictive power, when each individual predictor has little or no predictive power by itself [263]. This is particularly the case if non-linear interactions exist among multiple predictors. Feature selection techniques for machine learning [263], especially non-linear ones, can partially address this approach’s shortcoming. For example, we can use decision trees to identify important predictors and drop those predictors not in the decision trees.

3.10.4 Using physician practice predictors

With rare exceptions [264], existing clinical predictive models [38] use only patient predictors by assuming that a patient’s outcome depends only on the patient’s characteristics. In reality, a patient’s outcome depends not only on the patient’s characteristics but also on the treating physician’s practice characteristics. Based on this observation, we recently proposed using physician practice predictors to improve the prediction accuracy of various factors such as patient health outcomes [257, 261]. A physician’s practice profile includes his/her own (e.g., demographic) information, historical data of all his/her patients in the provider’s electronic medical record system and administrative systems, and other aspects of his/her practice that can influence an outcome. An example physician practice predictor is the indicator of whether the physician and the patient are of the same race or speak the same language. Another example physician practice predictor is the average outcome measure of a physician’s patients with a particular health issue. A third example physician practice predictor is the availability of clinic visit hours on weekends, which influences the risk for an ED visit in bronchiolitis patients. It could be beneficial to add physician practice predictors into the predictive models for bronchiolitis.

3.10.5 Using treatment model profile predictors

Besides the patient’s characteristics and the treating physician’s characteristics, the treatment model applied to the patient also affects the patient’s outcome [257]. Based on this observation, we recently proposed using treatment model profile
predictors to improve the prediction accuracy of various factors such as patient health outcomes [257]. It could be beneficial to add treatment model profile predictors into the predictive models for bronchiolitis.

3.10.6 Using large data sets and exhaustive variable sets

With rare exceptions such as Mansbach et al. [69], almost all existing work on predictive modeling for bronchiolitis has been conducted on small data sets including (typically much) fewer than 1,500 bronchiolitis patients. In general, a predictive model’s accuracy improves as the training data set becomes larger. This is particularly the case if the predictive model uses many predictors. By using data of more bronchiolitis patients to train the predictive models for bronchiolitis, we are likely to improve the predictive models’ accuracy.

An existing predictive model for bronchiolitis typically uses few variables. By using an exhaustive set of variables coupled with a large number of bronchiolitis patients, we are likely to further improve the predictive models’ accuracy.

4. Discussion

4.1 Main findings

Substantial effort has been invested in predictive modeling for bronchiolitis. Although considerable progress has been made, much remains to be done. From an algorithmic perspective, three types of research opportunities exist in predictive modeling for bronchiolitis. First, for many issues related to bronchiolitis, existing predictive models have inadequate accuracy. In this case, the research opportunity lies in improving the prediction accuracy. Second, for certain issues related to bronchiolitis, existing predictive models do not cover some parts of their full scope, such as a particular bronchiolitis patient population or a specific healthcare facility setting. The research opportunity here resides in constructing additional predictive models to cover their full scope. Third, for some issues related to bronchiolitis, no predictive model has ever been built. Here, the research opportunity is to develop and evaluate predictive models. Throughout the above presentation, we have identified multiple limitations and open problems in predictive modeling for bronchiolitis, and provided some preliminary thoughts on how to address them. This establishes a foundation for future research in this domain.

As is the case with clinical predictive modeling in general, to ensure that a predictive model for bronchiolitis has good performance, it will be important to test the predictive model on data sets from different institutions [24]. If the predictive model is developed on one institution’s data set and turns out to not work well on another institution’s data set, we can use domain adaptation techniques [265] and/or transfer learning techniques [269, 270] in machine learning to improve the predictive model’s generalizability.
The existing predictive models for bronchiolitis were mainly developed and tested on data from the same years. As various factors such as patient population characteristics may evolve over time, it would be essential to test the predictive models on data from subsequent years. This can help us understand how well the predictive models perform in practice.

So far, the research work on predictive modeling for bronchiolitis has focused on prediction accuracy. Nobody has ever translated a predictive model for bronchiolitis into routine clinical use and then measured the resulting improvement in patient outcome. This is another interesting area that deserves future research.

As mentioned in Wennberg et al. [266, 267], significant practice variation exists in managing many medical conditions rather than is unique to managing bronchiolitis. For instance, due to physicians’ lack of consensus on questions of safety and efficacy, admission rates for more than 80% of medical conditions and the rates for most operations are highly variable [268]. We would expect that after proper extension, machine learning predictive modeling techniques to be developed for bronchiolitis can be adapted to facilitate standardizing care for many other medical conditions.

4.2 Limitations

This systematic review has several limitations. First, by excluding articles not written in English, we may have missed predictive models for bronchiolitis published in other languages. Second, by limiting the systematic review to PubMed, we may have missed relevant articles not indexed in PubMed. Third, although we did not find any study that deployed a predictive model for bronchiolitis in operational settings and then measured the resulting improvement in patient outcome, it is still possible that this has been done for some of the predictive models described in this paper after the study was published. Likewise, there may be other predictive models for bronchiolitis for which this has been done, but those predictive models have never been published in peer-reviewed forums.

5. Conclusions

We systematically reviewed the literature on predictive modeling for bronchiolitis. Our results show that many problems remain open in predictive modeling for bronchiolitis. Future studies will need to address them to achieve optimal predictive models.

Authors’ contributions

GL had main responsibility for the manuscript. He conceptualized the presentation approach, conducted literature search and review, and drafted the manuscript. FN, PG, TG, and BS participated in conceptualizing the presentation approach, provided
feedback on various medical issues, and revised the manuscript. BS and FN also participated in literature search. All authors read and approved the final manuscript.

Acknowledgments

We thank Katherine Sward, Wendy W. Chapman, and Julio C. Facelli for helpful discussions. This work was supported in part by grant K08 HS018538 from the Agency for Healthcare Research and Quality.

Conflict of interest statement

The authors report no conflicts of interest.

Summary points

What is already known on the topic:

- Bronchiolitis is prevalent in young children. Significant practice variation exists among hospitals and physicians caring for children with bronchiolitis.
- To help standardize care for bronchiolitis, many predictive models have been proposed in the literature on various issues related to bronchiolitis.

What this study added to our knowledge:

- We provide a comprehensive review of the existing state of the art of predictive modeling for bronchiolitis.
- We identify several limitations and open problems in predictive modeling for bronchiolitis.
- We provide some preliminary thoughts on how to address these limitations and open problems. This establishes a foundation for future research in this domain.

References


43


Appendix: List of acronyms

- **AUC**: area under the receiver operating characteristic curve
- **BiPAP**: bilevel positive airway pressure
- **CART**: Classification And Regression Tree
- **CPAP**: continuous positive airway pressure
- **ED**: emergency department
- **HFNC**: high-flow nasal cannula
- **ICU**: intensive care unit
- **NIPPV**: non-invasive positive pressure ventilation
- **PCR**: polymerase chain reaction
- **RSV**: respiratory syncytial virus