

# Predicting Appropriate Admission of Bronchiolitis Patients in the Emergency Room: Rationale and Methods

Gang Luo<sup>1</sup>, PhD; Bryan L Stone<sup>2</sup>, MD, MS; Michael D Johnson<sup>2</sup>, MD; Flory L Nkoy<sup>2</sup>, MD, MS, MPH

<sup>1</sup>Department of Biomedical Informatics, University of Utah, Suite 140, 421 Wakara Way, Salt Lake City, UT 84108, USA

<sup>2</sup>Department of Pediatrics, University of Utah, 100 N Mario Capecchi Drive, Salt Lake City, UT 84113, USA

gang.luo@utah.edu, bryan.stone@hsc.utah.edu, mike.johnson@hsc.utah.edu, flory.nkoy@hsc.utah.edu

## Corresponding author:

Gang Luo, PhD

Department of Biomedical Informatics, University of Utah

Suite 140, 421 Wakara Way

Salt Lake City, UT 84108

USA

Phone: 1-801-213-3565

Fax: 1-801-581-4297

Email: gang.luo@utah.edu

## Abstract

**Background:** In young children, bronchiolitis is the most common illness resulting in hospitalization. For children less than age two, bronchiolitis incurs an annual total inpatient cost of \$1.73 billion. Each year in the U.S., 287,000 emergency department (ED) visits occur because of bronchiolitis, with a hospital admission rate of 32-40%. Due to a lack of evidence and objective criteria for managing bronchiolitis, ED disposition decisions (hospital admission or discharge to home) are often made subjectively resulting in significant practice variation. Studies reviewing admission need suggest that up to 29% of admissions from the ED are unnecessary. About 6% of ED discharges for bronchiolitis result in ED returns with admission. These inappropriate dispositions waste limited healthcare resources, increase patient and parental distress, expose patients to iatrogenic risks, and worsen outcomes.

Existing clinical guidelines for bronchiolitis offer limited improvement in patient outcomes. Methodological shortcomings include that the guidelines provide no specific thresholds for ED decisions to admit or to discharge, have an insufficient level of detail, and do not account for differences in patient and illness characteristics including co-morbidities. Predictive models are frequently used to complement clinical guidelines, reduce practice variation, and improve clinicians' decision making. Used in real time, predictive models can present objective criteria supported by historical data for an individualized disease management plan and guide admission decisions. However, existing predictive models for ED patients with bronchiolitis have limitations, including low accuracy and the assumption that the actual ED disposition decision was appropriate. To date, no operational definition of appropriate admission exists. No model has been built based on appropriate admissions, which include both actual admissions that were necessary and actual ED discharges that were unsafe.

**Objective:** The goal of this study is to develop a predictive model to guide appropriate hospital admission for ED patients with bronchiolitis.

**Methods:** This study will: (1) Develop an operational definition of appropriate hospital admission for ED patients with bronchiolitis. (2) Develop and test the accuracy of a new model to predict appropriate hospital admission for an ED patient with bronchiolitis. (3) Conduct simulations to estimate the impact of using the model on bronchiolitis outcomes.

**Results:** We are currently extracting administrative and clinical data from the enterprise data warehouse of an integrated healthcare system. Our goal is to finish this study by the end of 2019.

**Conclusions:** This study will produce a new predictive model that can be operationalized to guide and improve disposition decisions for ED patients with bronchiolitis. Broad use of the model would reduce iatrogenic risk, patient and parental distress, healthcare use, and costs and improve outcomes for bronchiolitis patients.

**Keywords:** Decision support techniques; forecasting; computer simulation; machine learning

## 1. Introduction

### The need for improving emergency department disposition decision making for bronchiolitis patients

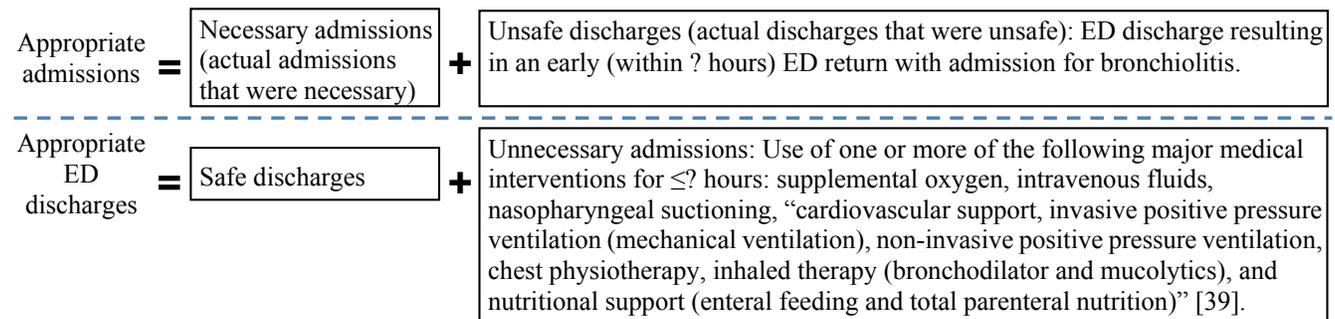
Bronchiolitis is inflammation of the bronchioles, the smallest air passages in the lungs, primarily seen in children less than age two. Within the first year of life, 10% of children are diagnosed with bronchiolitis [1]. By age two, >1/3 of children have had a bronchiolitis diagnosis [2]. Bronchiolitis causes about 71 hospitalizations and 77 emergency department (ED) visits per 1,000 infant years [3]. In the U.S., each year bronchiolitis incurs around 287,000 ED visits [4], 128,000 hospitalizations [5], and

\$1.73 billion in total inpatient costs (2009) [5]. For children under age two, bronchiolitis is the most common cause of hospitalization and represents 16% of all hospitalizations [5-8].

Despite the huge burden of bronchiolitis care, hospitalization decisions are made with insufficient evidence [7, 9] resulting in variable admission rates [1, 6, 9-19]. 32-40% of ED patients with bronchiolitis are admitted to the hospital [20-22]. Studies suggest that 20-29% of these admissions are unnecessary [23, 24]. Unnecessary admissions waste healthcare resources, overwhelm hospital capacity, increase patient and parental distress, introduce iatrogenic risk such as exposure to other infectious diseases, and expose other hospitalized children to the respiratory pathogens of these patients [11, 17, 25]. As many as 10% of infants affected by bronchiolitis have adverse events while in the hospital [26]. Similarly, about 6% of ED discharges for bronchiolitis are unsafe, resulting in ED return with hospital admission [27] due to inadequate treatment [11]. New approaches are needed to improve ED disposition decision making and reduce unnecessary admissions and unsafe ED discharges.

One method to reduce practice variation and improve clinicians' decision making for bronchiolitis care is to use clinical guidelines [28-31]. However, existing clinical guidelines for bronchiolitis acknowledge that decisions to admit or to discharge are subjective and rely on variable clinical judgment due to a lack of specific objective thresholds with good evidence [30, 31]. Clinical guidelines also do not account for differences in patient and illness characteristics including comorbidities [32], and offer limited improvement in determining ED disposition.

Predictive models are an alternative method to reduce practice variation and improve decision making [20-22, 33-35]. Predictive models can present objective criteria supported by historical data for an individualized disease management plan. Using data from previous patient encounters to identify patterns, predictive models [36] can overcome the limitations of clinical guidelines. Predictive models can be incorporated into electronic decision support tools [37] to support the provisional judgment of clinicians or to trigger clinicians to reconsider their judgment in real time [20]. This is especially useful for physicians who see children infrequently or are junior. Typically when results of predictive models are provided, human experts can make better decisions [38, page 6].



**Figure 1.** The definition framework of appropriate admission vs. appropriate ED discharge that was provided in our paper [39]. The details denoted by ? will be determined by direct evidence in this study.

As reported in our paper [39], existing predictive models for ED patients with bronchiolitis are limited by low accuracy and the assumption that actual ED disposition decisions were appropriate. No operational definition of appropriate admission of ED patients with bronchiolitis exists and no model has been built on appropriate admissions (see Fig. 1). To fill the gap, we will: (1) develop an operational definition of appropriate hospital admission for bronchiolitis, (2) develop a new model to accurately predict appropriate disposition for ED patients with bronchiolitis, and (3) conduct simulations to estimate the impact of using the model on outcomes.

**Innovations**

This study makes the following innovations within the context of bronchiolitis:

- (1) We will develop a new approach to construct an operational definition of appropriate hospital admission in the ED based on objective data rather than clinical judgment. No such approach currently exists.
- (2) We will figure out the most important attributes to put into the predictive model using a new simulation method. We will use various attribute combinations to ascertain the minimum requirement on performance and permit tradeoffs for the adaption of our model further than our setting dependent upon available attributes. Current models are not generalizable beyond the study site because they rely on a certain set of attributes that may be non-existent in different electronic medical records.

- (3) We will build the first model to accurately predict appropriate admission for ED patients with bronchiolitis in real time. No such model currently exists. We will transform bronchiolitis care by developing a predictive model to guide appropriate admission for the first time.
- (4) Our model will increase prediction accuracy by using a rich set of extracted attributes, including known predictors of hospital admission not used in existing models for ED patients with bronchiolitis.
- (5) Our model will include environmental variables with a potential for further increase in accuracy. Air quality environmental variables are associated with the daily number of hospitalizations for bronchiolitis [40] and a child's risk of hospitalization for bronchiolitis within the first year of life [41]. The predictive power of air quality and respiratory virus environmental variables for appropriate admission has never been evaluated.
- (6) Our study will evaluate the impact of using the model on outcomes. Previous predictive models focused only on accuracy. No impact estimate of using the model on bronchiolitis outcomes has ever been provided.
- (7) We will use a large data set of 26,701 bronchiolitis patients with high potential to achieve high prediction accuracy. Previous studies are limited by small data sets with typically far fewer than 1,000 patients. Many useful predictors of hospital admission cannot be identified in small data sets.

In summary, this study is significant as it will fill gaps by developing a new model to guide and improve disposition decisions for ED patients with bronchiolitis. Broad use of the model will reduce iatrogenic risk, patient and parental distress, healthcare use, and cost and improve clinical outcomes for bronchiolitis patients. Future study will test the impact of using the model in a randomized controlled trial after implementing the model into an existing electronic medical record to facilitate real-time decision making.

## 2. Methods

Machine learning is a field that studies the automatic improvement of computer algorithms with experience. Machine learning methods, such as support vector machine, neural network, and decision tree, are commonly used in predictive modeling [36] and will be adopted in our study. In comparison to statistical methods, machine learning can improve prediction accuracy, occasionally doubling it, with less stringent assumptions on data distribution [38, 42, 43].

We will use the same patient population and data sets for all three aims:

- (1) **Patient population:** Our study cohort includes children under age two who had ED encounters at 22 Intermountain Healthcare facilities for bronchiolitis (ICD-9-CM discharge diagnosis code 466.1 [4]) in the past 10 years. In Utah, Intermountain Healthcare is the biggest healthcare system that has 185 clinics and 22 hospitals.
- (2) **Data sets:** A large administrative and clinical data set in the enterprise data warehouse (EDW) of Intermountain Healthcare will be used. The Intermountain Healthcare EDW contains a vast set of attributes [44]. Our Intermountain Healthcare data analyst will run SQL queries to obtain a data set that has been de-identified and encrypted, and then securely transfer it to a computer that is encrypted and password-protected. Secondary analysis will be conducted on the computer. Intermountain Healthcare has dedicated tables to identify changes in procedure and diagnosis codes. The data set contains electronic documentation of about 85% of pediatric care delivered in Utah [45] and includes approximately 400 attributes. A partial list of categories of these attributes includes: "admission date and time; age; orders (medications, labs, exams, immunizations, imaging, counseling, etc.), including order name, ordering provider, performing date, and result date; allergies;" chief complaint; "diagnoses; discharge date; exam result; facility seen for the patient visit; gender; health insurance; healthcare cost (billed charge, Intermountain Healthcare internal cost, and reimbursed cost); height; home address; immunizations; lab test result; language(s) spoken; medication refills; primary care physician as listed in the electronic medical record; problem list; procedure date; procedures; provider involved in the visit; race/ethnicity; referrals; religion; visit type (inpatient, outpatient, urgent care, or emergency department); vital signs; weight; ..." [46]

For the last five years, data captured cover more than 2,900 patients under age two and 3,500 ED encounters for bronchiolitis per year. Due to its attribute richness and large size, the data set provides many advantages in the exploration of the proposed predictive models. Furthermore, we will use 21 environmental variables that regional monitoring stations recorded over the past decade within the Intermountain Healthcare region. These variables include carbon monoxide (CO), nitrogen dioxide (NO<sub>2</sub>), particulate matter up to 2.5 μm in size (PM<sub>2.5</sub>) and 10 μm in size (PM<sub>10</sub>), ozone (O<sub>3</sub>), sulfur dioxide (SO<sub>2</sub>), relative humidity, temperature, precipitation, wind speed, dew point, and activities of each of the following viruses: enterovirus, adenovirus, parainfluenza virus types 1, 2, and 3, human metapneumovirus, influenza A virus, influenza B virus, rhinovirus, and respiratory syncytial virus. The data for all non-virus environmental variables came from federal data sources [47, 48], which provide such data throughout the United States. Observation unit admissions will be treated as hospital admissions, as the one pediatric observation unit within Intermountain Healthcare has the same admission, coding, billing, and documentation requirements. Our analysis will consider various attribute combinations to ascertain the minimum requirement on performance and permit tradeoffs for the adaption of our model further than our setting dependent upon available attributes. Our analysis

results will serve as the basis for expansion of our models to other clinical data sets and diseases beyond bronchiolitis in the future.

**Aim 1: Develop an operational definition of appropriate hospital admission for ED patients with bronchiolitis.**

In a recent paper [39], we provided a definition framework of appropriate hospital admissions. As shown in Fig. 1, appropriate admissions = necessary admissions + unsafe discharges. Appropriate ED discharges = safe discharges + unnecessary admissions. The definition uses several threshold values, such as the maximum number of hours for which major medical interventions are used. Using a data-driven approach, we will fill in these values and develop an operational definition to be used in Aims 2 and 3.

For unsafe discharges, we will examine the distribution of the interval between discharge from the ED and a return visit resulting in admission for bronchiolitis within the period of two weeks [49, 50]. The 95<sup>th</sup> percentile of the interval will cover most readmissions and define the return threshold for unsafe discharge. The distribution is highly skewed toward a short interval [27]. Thus, the return threshold will be insensitive to the length of the period chosen.

For unnecessary admissions, we will examine the patients who stayed in the hospital for  $\leq 12$  hours and were discharged without readmission for bronchiolitis within two weeks. These patients are likely to have been admitted unnecessarily. Their median duration of using major medical interventions (Fig. 1) will serve as a conservative threshold for use of major medical interventions in all admissions. Unnecessary admissions are those with major medical intervention exposures for  $\leq$  the threshold. We will conduct sensitivity analysis to evaluate the impact of interactions between major medical interventions and other variables.

If the operational definition for all bronchiolitis patients lacks face validity, we will examine data distributions for different age groups to obtain operational definitions by age group. Since the medical interventions for bronchiolitis have not changed over the last 10 years, we would expect the operational definition to remain the same during this period.

**Aim 2: Develop and test the accuracy of a new model to predict appropriate hospital admission for an ED patient with bronchiolitis.**

We will use clinical, administrative, and environmental variable attributes to build machine learning models to predict appropriate hospital admission for individual ED patients with bronchiolitis.

**Data pre-processing:** Traditional techniques like imputation will be used to handle missing values and identify and correct/remove invalid values [36, 51]. In the case of environmental variables, classic methods [40, 41] will be used to extract aggregate values (e.g., daily average) from raw values. In the case of clinical and administrative attributes, grouper models like the DCG (Diagnostic Cost Group) system will be used to aggregate diseases, drugs, and procedures to reduce attributes [52, Chapter 5].

**Input variables** (a.k.a. independent variables): For ED patients with bronchiolitis, predictors of hospital admission have not been exhaustively identified. We compiled in our recent paper [39] a comprehensive list of known predictors. Some of these known predictors (e.g., atopic dermatitis [53], low dew point [54], duration of respiratory distress [7], absence of familial atopy [55], enterovirus infection [55], etc.) have not been used in existing predictive models for ED patients with bronchiolitis. All known predictors stored in the Intermountain Healthcare EDW and environmental data sets will be used as input variables. In addition, our data sets contain attributes beyond known predictors. We will use classic feature selection techniques [56] like the information gain method to find attributes likely to be predictive of appropriate admission. Our team's clinical experts will review attributes, select attributes with face validity, and add these as input variables. With more new predictors of appropriate hospital admission and larger sample size, we anticipate higher prediction accuracy.

**Predictive models:** We will use Weka [56] to construct predictive models. Weka is a widely used open-source machine learning toolkit. It integrates a large set of standard machine learning algorithms and feature selection techniques. Both categorical and numerical variables exist in administrative, clinical, and environmental data. Supervised machine learning algorithms that can deal with both categorical and numerical variables, such as  $k$ -nearest neighbor and random forest, will be used. We will examine each applicable algorithm and tune hyper-parameters manually.

The classic area under the receiver operating characteristic curve (AUC) [56] performance metric will be used. Our target will be models achieving an  $AUC \geq 0.9$ , which is considered outstanding discrimination [57, page 177]. Some machine learning models, such as decision tree and  $k$ -nearest neighbor (similar patients), can be more easily interpreted [58, 59]. Other machine learning models, such as random forest, are less straightforward to interpret. If accuracies of models are comparable (an  $AUC \geq 0.9$  and  $\leq 0.02$  worse for interpretable models compared to less interpretable models), we will favor those that clinicians can more easily interpret.

**Sample size justification and performance evaluation:** We have 10 years of data. We will train and test predictive models using a standard method. We will perform stratified 10-fold cross validation [56, Section 5.3] on the initial nine years of data to train predictive models and provide estimates of their accuracy. Data from the tenth year will be used to evaluate performance of the best-performing machine learning algorithm, reflecting use in practice. To figure out the environmental variable,

administrative, and clinical attributes necessary for high accuracy, we will use backward elimination [36] to remove input variables so long as the AUC does not decrease by more than 0.02 or below 0.90.

No AUC achieved by current care has been reported before. By extrapolating from statistics reported in the literature (unnecessary admissions up to 29% and unsafe ED discharges of 6%), we anticipate the AUC achieved by current care to be between 0.6 and 0.8+ [20-24, 27]. We will test the hypothesis that the model’s prediction will be more accurate by a difference in AUC of  $\geq 0.05$ . The dependent variable has two possible values: appropriate hospital admission and appropriate ED discharge. Assuming a correlation coefficient of 0.6 between the model’s prediction result and the actual disposition decision for both values and using a one-sided Z-test at a 0.05 significance level, a sample size of 356 instances per possible value of the dependent variable will have 90% power to detect an AUC increase of 0.05. Data from the tenth year include 3,615 ED visits for bronchiolitis, which provides adequate power for testing our hypothesis.

Based on two prior studies’ results, we anticipate that our model will achieve an AUC  $\geq 0.9$  and outperform current care in making disposition decisions. Neither prior study on predicting a bronchiolitis patient’s ED disposition is similar to our study that uses appropriate admission as the gold standard. The first study [21] used actual admission as the gold standard and achieved an AUC of 0.87. The second study [34] used judgment of an attending pediatrician as well as a length of stay  $>1$  day as the gold standard. The predictive model achieved 81% accuracy, better than an average admitting resident’s disposition decision.

For ED patients with bronchiolitis, 17 known predictors of hospital admission (listed in Table 1) are consistently recorded at Intermountain Healthcare facilities and available as structured attributes in our data sets, along with many other potential predictors. We will start building our model using structured attributes. If the model cannot achieve high prediction accuracy, we will extract additional de-identified input variables from ED clinical notes by conducting medical natural language processing on the HIPAA-compliant Homer computer cluster at the University of Utah [60]. For instance, additional input variables include the 8 known predictors of hospital admission (listed in Table 1) that are inconsistently recorded in clinical notes at Intermountain Healthcare facilities.

**Table 1.** The list of known predictors of hospital admission for ED patients with bronchiolitis recorded at Intermountain Healthcare facilities.

Category	Predictors
The known predictors that are consistently recorded at Intermountain Healthcare facilities and available as structured attributes in our data sets	SpO <sub>2</sub> , heart rate, respiratory rate, temperature, age, gender, prior hospitalization, prior intubation, abnormal chest x-ray, low dew point (from the environmental variable data set), rhinovirus infection, co-infection, dehydration, history of bronchopulmonary dysplasia, history of eczema, prematurity, maternal/passive smoking
The known predictor that is rarely recorded as structured attributes at Intermountain Healthcare facilities	enterovirus infection
The known predictors that are inconsistently recorded in clinical notes at Intermountain Healthcare facilities	increased work of breathing, poor feedings, decreased feeding, breastfed, abnormalities on auscultation, retractions, family history of atopy, fewer albuterol in the first hour

If our model still cannot reach high prediction accuracy on the entire group of ED patients with bronchiolitis, we will conduct subanalyses to identify subgroups of ED patients with bronchiolitis on which our model performs well. In this scenario, we will apply our final model only to the identified subgroups of patients. These subgroups are identified by certain characteristics, such as co-morbidity, prematurity, age, or ED arrival time (daytime vs. night, weekday vs. weekend, etc.) that are typically independent variables in the original model.

We have large data sets. If scalability is a problem with Weka, a parallel machine learning toolkit like Spark’s MLlib [61] will be adopted to develop predictive models on the secure Homer computer cluster [60].

**Aim 3: Conduct simulations to estimate the impact of using the model on bronchiolitis outcomes.**

We will use a method similar to that in Luo *et al.* [46] to establish the model’s utility for future use in clinical practice. More specifically, we will estimate the impact of using the model on bronchiolitis outcomes by applying the model to a retrospective cohort, and determine how the model can be generalized to different sites that collect differing sets of attributes. Our model will be developed on data from Intermountain Healthcare. Our simulations will help determine how to implement the model in other EDs. No prior study has either assessed the impact of using a predictive model on bronchiolitis outcomes or found the set of attributes most essential to generalize the model.

**Outcomes:** We will assess the outcomes of hospital admission, discharge to home, cost, and ED return. The primary outcome is cost. Other outcomes are indirectly reflected in cost and secondary. Each medical claim is companioned by a billed cost, a reimbursed cost, and an Intermountain Healthcare internal cost [52, page 43]. The Intermountain Healthcare internal cost [62] will be used because it is subject to less variation resulting from member cost-sharing [52, page 45] and more closely reflects actual cost. To deal with inflation, the medical consumer price index [63] will be used to standardize costs to 2014 US dollars. ED returns will be computed using the time interval defining unsafe discharge.

**Estimate a model's impact:** Given a predictive model and a set of input variables, we will estimate the impact of using the model on each outcome. The same method in aim 2 will be used to train the model on data from the first nine years. Data from the tenth year have four groups (Fig. 1): necessary and unnecessary admissions, and unsafe and safe discharges. For each group, we will obtain prediction results, then estimate the outcome if the model's suggestions were followed. For example, consider necessary admissions. The model will erroneously predict that some of these patients should be discharged. We assume that in clinical application, every such patient will incur an unsafe discharge, an early return visit for bronchiolitis, and a cost equal to unsafe discharges' average cost. The overall estimated outcome is the aggregate of outcome estimates in all four groups. Similarly, we can determine the minimum requirement of the model's accuracy for the model to be valuable clinically.

**Sensitivity analysis:** Intermountain Healthcare gathers a vast range of attributes. A different hospital may gather a portion of these attributes. To assure that the model is generalizable, we will examine miscellaneous attribute combinations and estimate the outcomes of bronchiolitis when using the modified model. Our estimate will determine which attributes are important to include. In the case that an important attribute is non-existent in a given ED, the estimate can advise substitute attributes that have a minor impact on bronchiolitis outcomes.

Our complete model will include as many as 400 attributes. Conducting simulations for each possible combination of the attributes is not realistic because of the exponential growth of the number of combinations. As an alternative, an attribute grouping approach will be used. This approach associates attributes that commonly co-exist based upon the judgment of our clinical experts. If an attribute in a group is not recorded by a hospital, related attributes in the group are also likely to be missing, such as attributes from the same lab test panel. Grouping will allow us to create and publish a table that lists the groups of possible attribute combinations, including bronchiolitis outcomes estimated via simulations and the trained parameters of the predictive model. If a hospital shows interest in implementing the model, the table can help assess expected outcomes in their environment, whether additional attributes need to be gathered, and if so, which ones. One row in the table will reflect the attributes in the PHIS+ [64] data model that standardizes administrative and clinical attributes from six major U.S. children's hospitals. The model of the row will apply directly to at least these six hospitals.

**Sample size justification and performance evaluation:** We will test three hypotheses: use of our predictive model will be linked to reduced (1) costs, (2) ED returns, and (3) hospital admissions. Due to their skewed distribution, cost data will be log-transformed [52, page 134]. The primary hypothesis will be accepted if the model lowers the log cost by at least 10% of its standard deviation. We will use one-sided paired-sample t-test to assess the log cost difference between the model's prediction result and the actual disposition decision. We will use McNemar's test to assess the difference in ED returns and hospital admissions. A sample size of 857 data instances has 90% power to support the primary hypothesis at a 0.05 significance level. Data from the tenth year include 3,615 ED visits for bronchiolitis, which offers sufficient power for testing the primary hypothesis.

If it is too slow to perform simulations on one computer for the numerous combinations of attribute groups, we will conduct parallel simulations on the secure Homer computer cluster [60].

#### **Ethics approval**

We have already secured institutional review board approvals from Intermountain Healthcare and the University of Utah for this study.

### **3. Results**

At present, we are extracting administrative and clinical data from the EDW of Intermountain Healthcare. Our goal is to finish this study by the end of 2019.

### **4. Discussion**

The principle of our approach to developing an operational definition of appropriate hospital admission in the ED is general and can be used for other diseases beyond bronchiolitis. Our simulation method will ascertain how a predictive model can be

generalized to different sites collecting various sets of attributes, as well as the group of attributes most essential for generalization. This study will use data from a big healthcare system with numerous heterogeneous facilities spread across a large area. These facilities include EDs at 22 hospitals, ranging from community metropolitan and rural hospitals attended by general practitioners and family doctors with constrained pediatric resources to tertiary care children's and general hospitals in urban areas attended by sub-specialists. Each of these facilities has a differing patient population, scope of services, geographic location, staff composition, and cultural background. This variation creates a realistic situation for identifying factors that are generalizable to other facilities across the U.S. One of the models produced during simulation will directly apply to at least six large children's hospitals across the U.S. Moreover, this study will produce a new modeling strategy that can generalize to other clinical conditions where decision making is uncertain.

In summary, our work will transform bronchiolitis care by developing a new predictive model to guide appropriate admission for ED patients with bronchiolitis. Broad use of the model will lower healthcare use and cost and improve clinical outcomes for bronchiolitis patients. We will have a new simulation method to estimate the impact of using a predictive model on outcomes in dissimilar data environments. The method can be useful for implementing other models.

## **Acknowledgments**

We thank Maureen A. Murtaugh, Xiaoming Sheng, Michael Mundorff, Farrant Sakaguchi, Karen Valentine, Wendy W. Chapman, Adam B. Wilcox, Christopher G. Maloney, Chris Benitez, Bart Dodds, David E. Jones, Bret Heale, Katherine Sward, and John Holmen for helpful discussions.

## **Authors' contributions**

GL was mainly responsible for the paper. He performed literature review, formed the concept and developed the study, and wrote the paper. BS, MJ, and FN provided feedback on various medical issues, contributed to conceptualizing and designing the study, and revised the paper.

## **Conflicts of interest**

None declared.

## **References**

1. Ducharme FM. Management of acute bronchiolitis. *BMJ* 2011;342:d1658. PMID:21471173
2. Zorc JJ, Hall CB. Bronchiolitis: recent evidence on diagnosis and management. *Pediatrics* 2010;125(2):342-9. PMID:20100768
3. Carroll KN, Gebretsadik T, Griffin MR, Wu P, Dupont WD, Mitchel EF, Enriquez R, Hartert TV. Increasing burden and risk factors for bronchiolitis-related medical visits in infants enrolled in a state health care insurance plan. *Pediatrics* 2008;122(1):58-64. PMID:18595987
4. Hasegawa K, Tsugawa Y, Brown DF, Mansbach JM, Camargo CA Jr. Temporal trends in emergency department visits for bronchiolitis in the United States, 2006 to 2010. *Pediatr Infect Dis J* 2014;33(1):11-8. PMID:23934206
5. Hasegawa K, Tsugawa Y, Brown DF, Mansbach JM, Camargo CA Jr. Trends in bronchiolitis hospitalizations in the United States, 2000-2009. *Pediatrics* 2013;132(1):28-36. PMID:23733801
6. Mansbach JM, Emond JA, Camargo CA Jr. Bronchiolitis in US emergency departments 1992 to 2000: epidemiology and practice variation. *Pediatr Emerg Care* 2005;21(4):242-7. PMID:15824683
7. Parker MJ, Allen U, Stephens D, Lalani A, Schuh S. Predictors of major intervention in infants with bronchiolitis. *Pediatr Pulmonol* 2009;44(4):358-63. PMID:19283838
8. Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, Anderson LJ. Bronchiolitis-associated hospitalizations among US children, 1980-1996. *JAMA* 1999;282(15):1440-6. PMID:10535434
9. Brand PL, Vaessen-Verberne AA. Differences in management of bronchiolitis between hospitals in the Netherlands. *Eur J Pediatr* 2000;159(5):343-7. PMID:10834519
10. Behrendt CE, Decker MD, Burch DJ, Watson PH. International variation in the management of infants hospitalized with respiratory syncytial virus. *Eur J Pediatr* 1998;157(3):215-20. PMID:9537488
11. Chamberlain JM, Patel KM, Pollack MM. Association of emergency department care factors with admission and discharge decisions for pediatric patients. *J Pediatr* 2006;149(5):644-9. PMID:17095336
12. Johnson DW, Adair C, Brant R, Holmwood J, Mitchell I. Differences in admission rates of children with bronchiolitis by pediatric and general emergency departments. *Pediatrics* 2002;110(4):e49. PMID:12359822
13. Mallory MD, Shay DK, Garrett J, Bordley WC. Bronchiolitis management preferences and the influence of pulse oximetry and respiratory rate on the decision to admit. *Pediatrics* 2003;111(1):e45-51. PMID:12509594

14. Plint AC, Johnson DW, Wiebe N, Bulloch B, Pusic M, Joubert G, Pianosi P, Turner T, Thompson G, Klassen TP. Practice variation among pediatric emergency departments in the treatment of bronchiolitis. *Acad Emerg Med* 2004;11(4):353-60. PMID:15064208
15. Vogel AM, Lennon DR, Harding JE, Pinnock RE, Graham DA, Grimwood K, Pattemore PK. Variations in bronchiolitis management between five New Zealand hospitals: can we do better? *J Paediatr Child Health* 2003;39(1):40-5. PMID:12542811
16. Willson DF, Horn SD, Hendley JO, Smout R, Gassaway J. Effect of practice variation on resource utilization in infants hospitalized for viral lower respiratory illness. *Pediatrics* 2001;108(4):851-5. PMID:11581435
17. Willson DF, Jiao JH, Hendley JO, Donowitz L. Invasive monitoring in infants with respiratory syncytial virus infection. *J Pediatr* 1996;128(3):357-62. PMID:8774504
18. Wang EE, Law BJ, Boucher FD, Stephens D, Robinson JL, Dobson S, Langley JM, McDonald J, MacDonald NE, Mitchell I. Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) study of admission and management variation in patients hospitalized with respiratory syncytial viral lower respiratory tract infection. *J Pediatr* 1996;129(3):390-5. PMID:8804328
19. Christakis DA, Cowan CA, Garrison MM, Molteni R, Marcuse E, Zerr DM. Variation in inpatient diagnostic testing and management of bronchiolitis. *Pediatrics* 2005;115(4):878-84. PMID:15805359
20. Marlais M, Evans J, Abrahamson E. Clinical predictors of admission in infants with acute bronchiolitis. *Arch Dis Child* 2011;96(7):648-52. PMID:21339199
21. Laham FR, Trott AA, Bennett BL, Kozinetz CA, Jewell AM, Garofalo RP, Piedra PA. LDH concentration in nasal-wash fluid as a biochemical predictor of bronchiolitis severity. *Pediatrics* 2010;125(2):e225-33. PMID:20100751
22. Corneli HM, Zorc JJ, Holubkov R, Bregstein JS, Brown KM, Mahajan P, Kuppermann N; Bronchiolitis Study Group for the Pediatric Emergency Care Applied Research Network. Bronchiolitis: clinical characteristics associated with hospitalization and length of stay. *Pediatr Emerg Care* 2012;28(2):99-103. PMID:22270499
23. Shaw KN, Bell LM, Sherman NH. Outpatient assessment of infants with bronchiolitis. *Am J Dis Child* 1991;145(2):151-5. PMID:1994678
24. Shafik MH, Seoudi TM, Raway TS, Al Harbash NZ, Ahmad MM, Al Mutairi HF. Appropriateness of pediatric hospitalization in a general hospital in Kuwait. *Med Princ Pract* 2012;21(6):516-21. PMID:22678120
25. Mansbach JM, Clark S, Christopher NC, LoVecchio F, Kunz S, Acholonu U, Camargo CA Jr. Prospective multicenter study of bronchiolitis: predicting safe discharges from the emergency department. *Pediatrics* 2008;121(4):680-8. PMID:18381531
26. McBride SC, Chiang VW, Goldmann DA, Landrigan CP. Preventable adverse events in infants hospitalized with bronchiolitis. *Pediatrics* 2005;116(3):603-8. PMID:16140699
27. Norwood A, Mansbach JM, Clark S, Waseem M, Camargo CA Jr. Prospective multicenter study of bronchiolitis: predictors of an unscheduled visit after discharge from the emergency department. *Acad Emerg Med* 2010;17(4):376-82. PMID:20370776
28. Seiden JA, Scarfone RJ. Bronchiolitis: An evidence-based approach to management. *Clinical Pediatric Emergency Medicine* 2009;10(2):75-81. doi:10.1016/j.cpem.2009.03.006
29. Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadowski AM, Johnson DW, Light MJ, Maraqa NF, Mendonca EA, Phelan KJ, Zorc JJ, Stanko-Lopp D, Brown MA, Nathanson I, Rosenblum E, Sayles III S, Hernandez-Cancio S. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics* 2014;134:e1474-e1502. PMID:25349312
30. Scottish Intercollegiate Guidelines Network. Bronchiolitis in children. A national clinical guideline. <https://www.guidelinecentral.com/summaries/bronchiolitis-in-children-a-national-clinical-guideline/>. Archived at <http://www.webcitation.org/6bhQvWHzA>
31. Bronchiolitis Guideline Team, Cincinnati Children's Hospital Medical Center: Evidence-based care guideline for management of first time episode bronchiolitis in infants less than 1 year of age. <http://www.guideline.gov/content.aspx?id=34411>. Archived at <http://www.webcitation.org/6bhOyjITC>
32. Shaneyfelt TM, Centor RM. Reassessment of clinical practice guidelines: go gently into that good night. *JAMA* 2009;301(8):868-9. PMID:19244197
33. Walsh P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R. A validated clinical model to predict the need for admission and length of stay in children with acute bronchiolitis. *Eur J Emerg Med* 2004;11(5):265-72. PMID:15359199
34. Walsh P, Cunningham P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R. An artificial neural network ensemble to predict disposition and length of stay in children presenting with bronchiolitis. *Eur J Emerg Med* 2004;11(5):259-64. PMID:15359198
35. Destino L, Weisgerber MC, Soung P, Bakalarski D, Yan K, Rehborg R, Wagner DR, Gorelick MH, Simpson P. Validity of respiratory scores in bronchiolitis. *Hosp Pediatr* 2012;2(4):202-9. PMID:24313026

36. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York, NY: Springer; 2009. ISBN:038777243X
37. Greenes RA. *Clinical Decision Support: The Road to Broad Adoption*, 2nd ed. London, UK: Academic Press; 2014. ISBN:0123984769
38. Kuhn M, Johnson K. *Applied Predictive Modeling*. New York, NY: Springer; 2013. ISBN:1461468485
39. Luo G, Nkoy FL, Gesteland PH, Glasgow TS, Stone BL. A systematic review of predictive modeling for bronchiolitis. *Int J Med Inform* 2014;83(10):691-714. PMID: 25106933
40. Ségala C, Poizeau D, Mesbah M, Willems S, Maidenberg M. Winter air pollution and infant bronchiolitis in Paris. *Environ Res* 2008;106(1):96-100. PMID:17585899
41. Karr CJ, Demers PA, Koehoorn MW, Lencar CC, Tamburic L, Brauer M. Influence of ambient air pollutant sources on clinical encounters for infant bronchiolitis. *Am J Respir Crit Care Med* 2009;180(10):995-1001. PMID:19713450
42. Axelrod RC, Vogel D. Predictive modeling in health plans. *Disease Management & Health Outcomes* 2003;11(12):779-87. doi:10.2165/00115677-200311120-00003
43. Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. *PLoS One* 2014;9(2):e88225. PMID:24520356
44. Evans RS, Lloyd JF, Pierce LA. Clinical use of an enterprise data warehouse. *AMIA Annu Symp Proc* 2012:189-98. PMID:23304288
45. Byington CL, Reynolds CC, Korgenski K, Sheng X, Valentine KJ, Nelson RE, Daly JA, Osguthorpe RJ, James B, Savitz L, Pavia AT, Clark EB. Costs and infant outcomes after implementation of a care process model for febrile infants. *Pediatrics* 2012;130(1):e16-24. PMID:22732178
46. Luo G, Stone BL, Sakaguchi F, Sheng X, Murtaugh MA. Using computational approaches to improve risk-stratified patient management: rationale and methods. *JMIR Res Protoc* 2015;4(4):e128. PMID:26503357
47. Air quality data homepage of the United States Environmental Protection Agency. <http://www.epa.gov/airdata/>. Archived at <http://www.webcitation.org/6eJwV4Yn2>
48. MesoWest homepage. <http://mesowest.utah.edu/>. Archived at <http://www.webcitation.org/6eJwZa7gU>
49. Halfon P, Egli Y, van Melle G, Chevalier J, Wasserfallen JB, Burnand B. Measuring potentially avoidable hospital readmissions. *J Clin Epidemiol* 2002;55(6):573-87. PMID:12063099
50. Heggstad T, Lilleeng SE. Measuring readmissions: focus on the time factor. *Int J Qual Health Care* 2003;15(2):147-54. PMID:12705708
51. Pyle D. *Data Preparation for Data Mining*. San Francisco, CA: Morgan Kaufmann; 1999. ISBN:1558605290
52. Duncan I. *Healthcare Risk Adjustment and Predictive Modeling*. Winsted, CT: ACTEX Publications Inc.; 2011. ISBN:1566987695
53. Al-Shehri MA, Sadeq A, Quli K. Bronchiolitis in Abha, Southwest Saudi Arabia: viral etiology and predictors for hospital admission. *West Afr J Med* 2005;24(4):299-304. PMID:16483044
54. Wang VJ, Cavagnaro CS, Clark S, Camargo CA Jr, Mansbach JM. Altitude and environmental climate effects on bronchiolitis severity among children presenting to the emergency department. *J Environ Health* 2012;75(3):8-15. PMID:23091965
55. Nascimento MS, Souza AV, Ferreira AV, Rodrigues JC, Abramovici S, Silva Filho LV. High rate of viral identification and coinfections in infants with acute bronchiolitis. *Clinics (Sao Paulo)* 2010;65(11):1133-7. PMID:21243286
56. Witten IH, Frank E, Hall MA. *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Burlington, MA: Morgan Kaufmann; 2011. ISBN:0123748569
57. Hosmer Jr. DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*, 3rd ed. Hoboken, NJ: Wiley; 2013. ISBN:0470582472
58. Freitas AA. Comprehensible classification models: a position paper. *SIGKDD Explorations* 2013;15(1):1-10. doi:10.1145/2594473.2594475
59. Vellido A, Martín-Guerrero JD, Lisboa PJG. Making machine learning models interpretable. *Proc. ESANN* 2012: 163-72.
60. Bradford W, Hurdle JF, LaSalle B, Facelli JC. Development of a HIPAA-compliant environment for translational research data and analytics. *J Am Med Inform Assoc* 2014;21(1):185-9. PMID:23911553
61. MLlib homepage. <https://spark.apache.org/mllib/>. Archived at <http://www.webcitation.org/6agakLPsf>
62. James BC, Savitz LA. How Intermountain trimmed health care costs through robust quality improvement efforts. *Health Aff* 2011;30(6):1185-91. PMID:21596758
63. Consumer Price Index - Measuring price change for medical care in the CPI. <http://www.bls.gov/cpi/cpifact4.htm>. Archived at <http://www.webcitation.org/6agaajPCS>
64. Narus SP, Srivastava R, Gouripeddi R, Livne OE, Mo P, Bickel JP, de Regt D, Hales JW, Kirkendall E, Stepanek RL, Toth J, Keren R. Federating clinical data from six pediatric hospitals: process and initial results from the PHIS+ Consortium. *AMIA Annu Symp Proc* 2011;2011:994-1003. PMID:22195159