The Effects of Mechanical Ventilation on the Development of Acute Respiratory Distress Syndrome

by

Xiaoming Jia

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author_____
Department of Electrical Engineering and Computer Science
May 29, 2007

Certified by_____
Roger Mark
Distinguished Professor in Health Science and Technology
Thesis Supervisor

Accepted by_____
Arthur C. Smith
Chairman, Department Committee on Graduate Students

The Effects of Mechanical Ventilation on the Development of Acute Respiratory Distress Syndrome

by

Xiaoming Jia

Submitted to the Department of Electrical Engineering and Computer Science
on May 29, 2007, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Computer Science and Engineering

Abstract

Acute Respiratory Distress Syndrome (ARDS) is a severe lung illness characterized by inflammation and fluid accumulation in the respiratory system. Historically, ARDS and other forms of respiratory failure have been treated using mechanical ventilation to help maintain gas exchange in the lungs. However, clinical investigators are beginning to discover the adverse effects of mechanical ventilation if it is not applied properly. Specifically, excessive ventilator volumes and pressures may exacerbate existing lung injury and increase hospital mortality. Furthermore, aggressive ventilation may cause lung injury and trigger an inflammatory response that is characteristic of ARDS. These findings have alarmed the critical care community, and many studies have been conducted to find mechanical ventilator settings that reduce mortality in patients with ARDS. However, there have been no firm recommendations on the optimal settings for patients who require ventilator therapy for reasons apart from respiratory failure.

In this thesis, we retrospectively examine a large medical database (MIMIC-II) to study the relationship between mechanical ventilation and the development of ARDS. Specifically, our goals are to (1) find patients who did not have ARDS at the beginning of mechanical ventilation but who later developed the disease; (2) identify physiologic and ventilator-associated risk factors for ARDS; and (3) develop a text analysis algorithm to automatically extract clinical findings from radiology (chest x-ray) reports.

Our findings suggest that acute respiratory distress syndrome is a relatively common illness in patients who require mechanical ventilation in the ICU (152 of 789 without ARDS at the outset eventually developed the disease). High plateau pressure (odds ratio 1.5 per 6.3 cmH_2O, $p < 0.001$) is the most important ventilator-associated risk factor for the development of new ARDS. Physiologic risk factors include high weight, low blood pH, high lactate, pneumonia, and sepsis. Thus it may be possible to reduce the occurrence of ventilator-induced lung injuries with careful pressure management. However, a randomized prospective study is needed to support this hypothesis.

Thesis Supervisor: Roger Mark
Title: Distinguished Professor in Health Science and Technology

Acknowledgments

This thesis deserves the acknowledgement of many mentors, friends, and family.

I am most grateful for having Dr. Roger Mark as my thesis supervisor (as well as pre-medical advisor and teacher). His high expectations and practical advice have helped me achieve more than I could have imagined, and the high quality of this research could not have been possible without his guidance. Dr. Mark's passion for physiology has inspired many students, including myself, to pursue careers in medicine and biomedical research. For these reasons, it has been a true privilege to have him as my mentor.

I am indebted to two physician-scientists and colleagues at the Beth Israel Deaconess Medical Center (BIDMC), Dr. Atul Malhotra and Dr. Daniel Talmor. Their expertise in pulmonary physiology has been invaluable to the design and implementation of this study. I have learned a tremendous amount from them in the past year, and look forward to doing additional research with them in the years to come.

I owe special thanks to Dr. Mohammed Saeed, who introduced me to this project and to our colleagues at BIDMC. He was one of my first mentors in the Lab for Computational Physiology (LCP), and has always had high hopes for my academic and research success.

To my friends at LCP: Dewang Shavdia, Mauricio Villarroel, Ali Saeed, Tushar Parlikar, James Sun, and others; you have made this past year one of the most memorable at MIT. I have benefitted enormously from your knowledge and expertise in Linux, Java, Matlab, LaTex, as well as in Life.

I sincerely thank my family: my mother, who has given me so much; my father, who will be in my memory always; Frank, who has treated me like a brother; my grandmother, who always believed in me; and members of my extended family, who have always cared.

This research was made possible by grant R01 EB001659 from the National Institute of Biomedical Imaging and Bioengineering.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The lungs are a vital part of the body's respiratory system, responsible for acquiring oxygen and removing excess carbon dioxide from the bloodstream. These processes occur in the lung alveoli, microscopic air sacs that facilitate gas exchange, as shown in Figure 1.1. Pulmonary capillaries surround the alveoli, allowing gases to diffuse across the thin membrane separating blood in the capillaries from inspired air. In normal respiratory function, blood high in carbon dioxide and low in oxygen is delivered to the lungs, and oxygenated blood with lower carbon dioxide is returned to systemic circulation. This process helps to maintain normal bodily metabolism and homeostasis.

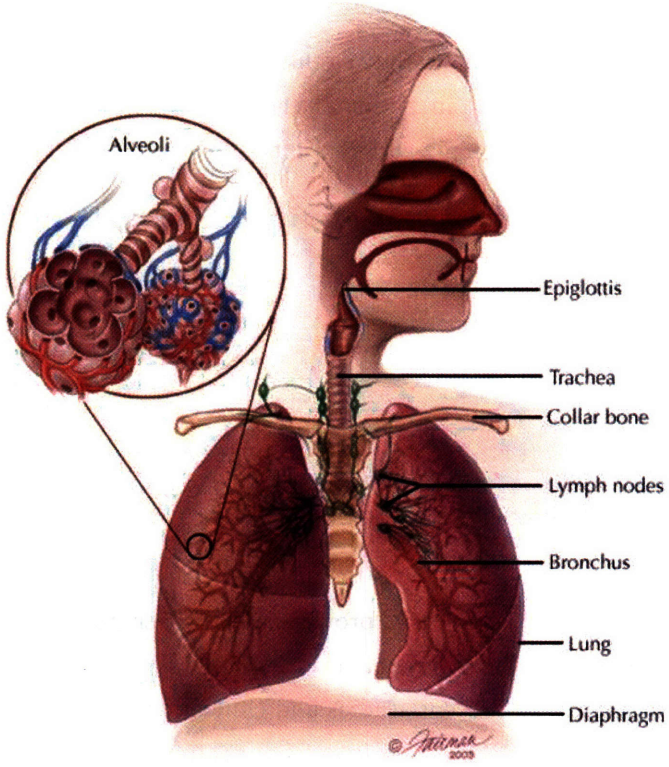


Figure 1.1: Anatomy of the respiratory system. Figure adapted from [1].

Failure of the respiratory system has negative consequences: 1) accumulation of carbon dioxide lowers blood pH and disrupts many biochemical activities; 2) inadequate tissue oxygenation may cause tissue death and organ failure, increasing the chance for patient death.

Acute respiratory distress syndrome (ARDS) is one of the most severe forms of respiratory failure. It is associated with prolonged hospitalizations and high mortality rates, making it a formidable complication to deal with in the intensive care unit (ICU). There is evidence that clinical interventions such as mechanical ventilation may influence the development of ARDS in patients at risk of the disease. Thus, there is tremendous value in identifying and understanding the risk factors for ARDS, especially if they are interventions that can be controlled in a clinical setting.

1.2 Thesis goals

The research conducted in this thesis aims to achieve the following goals:

- To retrospectively examine a large medical database (MIMIC-II) to find patients who do not have ARDS at the beginning of mechanical ventilation therapy, but who later develop the disease.
- To test the hypothesis that improper mechanical ventilation may cause acute lung injury (ALI)/acute respiratory distress syndrome (ARDS) in patients who do not have the diseases at the outset. We achieve this goal by identifying physiologic and ventilator-associated risk factors for ALI and ARDS.
- To design and evaluate an algorithm that automatically extracts clinical findings from chest x-ray reports.

1.3 Thesis outline

This thesis includes 5 chapters and 1 appendix.

- Chapter 2: *ARDS and Mechanical Ventilation*, provides a brief background on acute respiratory distress syndrome and mechanical ventilation. It explains the clinical criteria for ARDS, various ventilator settings and modes, and summarizes recent studies and clinical trials on ventilator-induced lung injury.
- Chapter 3: *Data Extraction and Statistical Methods*, describes the methods used to obtain and analyze data from the MIMIC-II database. It includes an overview of the database (how and from where the

data were collected, types of data available), the algorithms used for patient selection, and the basic theory of logistic regression and prediction.

- Chapter 4: *Results*, presents the important risk factors associated with the development of ARDS and the less severe form, acute lung injury (ALI). Physiologic and ventilator-associated risk factors are examined using univariate analyses, and their relative importance is compared using multivariate techniques.

- Chapter 5: *Discussion and Conclusions*, summarizes the important findings and discusses the results in the context of clinical relevance. Important milestones are listed, and suggestions for future work are also presented.

- Appendix A: *An Automated Radiology Report Reader*, presents a detailed description of the algorithm used to extract information from chest x-ray (text) reports. The algorithm evaluation is included, along with a summary of the Java source code.

Chapter 2

ARDS and Mechanical Ventilation

2.1 Acute Respiratory Distress Syndrome (ARDS)

Acute Respiratory Distress Syndrome is considered to be the leading cause of acute respiratory failure in the United States. It is a severe inflammatory disease that causes diffuse lung injury (accumulation of fluids and other blood contents) and impaired gas exchange in the alveoli. Other names for ARDS include wet lung, shock lung, leaky-capillary pulmonary edema, and adult respiratory distress syndrome. A milder form ARDS is called acute lung injury (ALI), which is a precursor to ARDS. In the United States, ARDS is responsible for 150,000 cases of respiratory failure per year and has an associated mortality rate of between 40% and 50%.

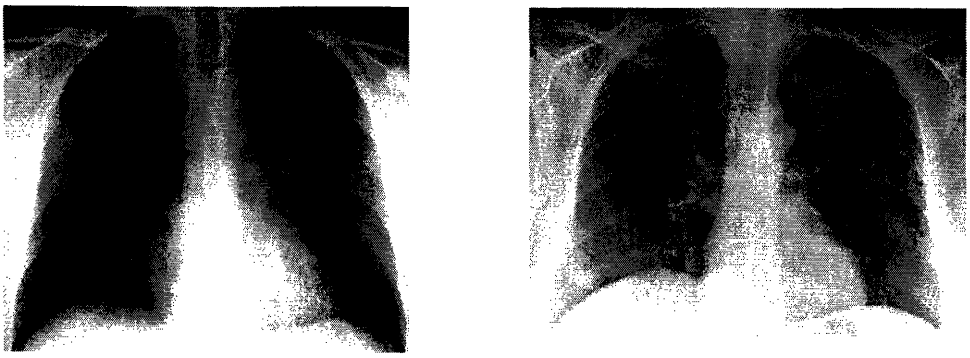
2.1.1 Clinical definition

The clinical definition for ALI and ARDS was established in 1994 by an American-European consensus conference [2] and includes the following criteria:

1. An acute onset.
2. Bilateral infiltrates revealed by a chest radiograph (x-ray).
3. Not left ventricular heart failure (pulmonary artery wedge pressure < 18 mmHg, or lack of evidence for heart failure).
4. (i) $PaO_2/FiO_2 < 300$ mmHg to be considered acute lung injury (ALI).
   (ii) $PaO_2/FiO_2 < 200$ mmHg to be considered acute respiratory distress syndrome (ARDS).

The consensus definition describes ARDS as having an acute or sudden onset, rather than a chronic progression. The chest x-ray must also show bilateral infiltrates (opaque or hazy regions in both the left and right lungs). An example of chest x-rays with bilateral infiltrates and clear lungs is shown in

Figure 2.1. When diagnosing ARDS, it is necessary to rule out the possibility of left ventricular heart failure, also known as congestive heart failure. In this illness, the heart is unable to pump out blood at an adequate rate, leading to high left ventricular filling pressures. The left atrial pressure also rises, causing increased pulmonary capillary hydrostatic pressure that forces fluid to enter lung alveoli. A standard method of distinguishing ARDS from cardiogenic pulmonary edema is to examine the pulmonary artery wedge pressure (PAWP), which reflects left arterial pressure: heart failure produces an elevated PAWP (over 18 mmHg) while ARDS does not. The fourth criteria examines the PaO_2/FiO_2 ratio, a measure of gas exchange in the lungs. PaO_2 is the partial pressure of oxygen in the blood, and FiO_2 is the fraction of oxygen in inspired air. Under normal conditions, PaO_2 is near 100 mmHg, FiO_2 is 0.21 (21% oxygen in free air), and the ratio PaO_2/FiO_2 is between 400 and 500 mmHg. In acute lung injury, PaO_2/FiO_2 drops below 300 mmHg while the more severe ARDS has a PaO_2/FiO_2 ratio below 200 mmHg. Such conditions describe "severe hypoxia refractory to oxygen," or low bodily oxygen content despite being treated with high amounts of oxygen [3].



(a) Normal lungs (b) Bilateral infiltrates

Figure 2.1: Comparison of chest x-rays in normal lungs vs. lungs with bilateral infiltrates, which is characteristic of ARDS. Infiltrates appear as opaque or hazy regions in the lungs.

2.1.2 Causes of ARDS

ARDS is triggered by a variety of direct and indirect injuries to the lungs. The most common causes are "inflammatory," in which systemic inflammation from another illness (such as trauma or sepsis) initiates a diffuse inflammatory injury in the lungs. Specifically, inflammatory mediators (cytokines and neutrophils) travel to the lungs via the bloodstream and cause pulmonary capillaries to become more permeable, allowing blood contents (fluid, cells, and proteins) to enter the lung alveoli. The presence of these infiltrates disrupt gas exchange and cause damage to the lung tissue.

A second class of insults is a result of physical injury, in which lung alveoli are damaged due to mechanical

stresses. For example, pulmonary contusion or lung overdistension due to mechanical ventilation may cause barotrauma (alveolar rupture) and injury of lung tissue, ultimately triggering an inflammatory response that leads to in lung infiltrates and disrupted gas exchange.

Lastly, there is ARDS due to infection. For example, pneumonia from a bacterial or viral infection may also cause a respiratory inflammatory response that results in infiltrate edema and impaired gas exchange. The causes of ARDS may vary, but it is possible that mechanical ventilation can influence one of more of these underlying illnesses.

2.1.3 Complications due to ARDS

The presence of ARDS often results in prolonged hospital stay and increased mortality due to various negative consequences. First, impaired gas exchange will cause hypoxia (inadequate oxygen levels) and hypercapnia (elevated carbon dioxide levels), both of which are life-threatening if left untreated. Second, it is possible for inflammation in the respiratory system to spread to other organs. The combined effects of systemic inflammation and hypoxia predisposes patients to multiple-organ failure, greatly increasing the chance of death. Third, diffuse infiltrates in the lungs may disturb the balance of surfactants in the alveoli, predisposing certain parts of the lungs to collapse. For this reason, it is possible for atelectasis and/or consolidation to be present at the same time as ARDS.

2.2 Mechanical ventilation

Mechanical ventilation is a clinical intervention used to assist or replace spontaneous breathing in patients for days to weeks in the intensive care unit. Its most important function is to maintain gas exchange in patients with respiratory failure (severe hypoxia and/or hypercapnia) or who cannot breathe on their own. In this thesis, "mechanical ventilation" refers to positive-pressure ventilation, where air is delivered to the lungs by applying positive pressures in the patient's airway. In order to control the delivered volumes and pressures, clinicians must perform intubation, a process by which endotracheal tube is passed through the mouth, the larynx, and into the trachea (Figure 2.2). In addition, the patient is usually sedated to prevent injurious interactions between spontaneous and mechanical breathing. For this reason, mechanical ventilation is considered an invasive intervention that has its own advantages and disadvantages.

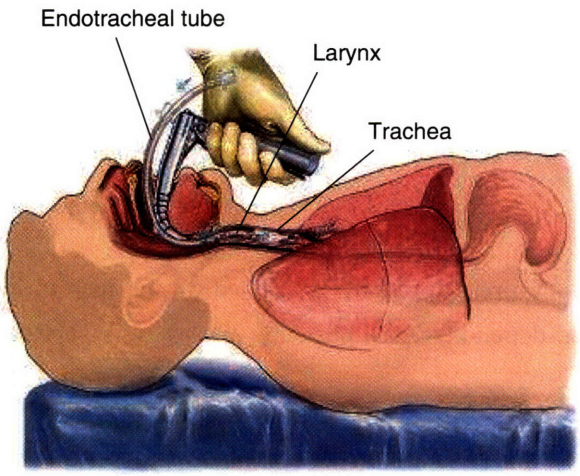


Figure 2.2: Endotracheal intubation for mechanical ventilation

2.2.1 Mechanical ventilator modes

A variety of ventilator modes are available to accommodate patients with different needs. There are two general categories of ventilator modes: volume-control and pressure-control. Volume-control modes deliver a fixed volume with each breath, while pressure-control modes apply a preset maximal pressure at the airway during inspiration to deliver breaths. Within each of the two categories, a variety of modes exist to accommodate different breathing patterns. Some modes deal with patients who are unable to breathe on their own, and others are for patients breathing spontaneously. The most common ventilator modes are discussed below and summarized in Table 2.1.

Volume-control modes:

- **Continuous Mandatory Ventilation (CMV)** - breaths are delivered at preset volumes and intervals regardless of patient effort. This mode is used only when a patient is sedated, paralyzed, or is apneic (not breathing) to minimize the chance of lung injury.
- **Assist Control Ventilation (A/C)** - the ventilator delivers a preset volume with each inspiratory effort. The inspiratory effort is detected by a drop in airway pressure as the patient begins to inhale. This mode prevents the ventilator from delivering a full tidal volume when the patient is maximally inhaled, a potential cause of barotrauma.
- **Intermittent Mandatory Ventilation (IMV)** - breaths are administered at a preset (lower) frequency, and the patient is allowed to breathe spontaneously between ventilator-delivered breaths. Syn-

chronous intermittent mandatory ventilation (SIMV) is similar to IMV, except that the ventilator-delivered breaths are administered according to patient inspiratory effort. IMV and SIMV may be used in "weaning," the process to help patients slowly breathe independently of the ventilator.

Pressure-control Modes:

- **Pressure Control (PC)** - a preset pressure is applied during inspiration at a fixed respiratory rate. The volume of air delivered depends on the patient's airway resistance, lung compliance, and duration of the inspiration period.
- **Pressure Support (PS)** - a preset amount of support pressure is used to assist every spontaneous breath. This mode differs from CMV and A/C in that the amount of pressure is set instead of the tidal volume. It has been recommended to use pressure support for patients who are breathing spontaneously but are still in need of assistance.
- **Continuous Positive Airway Pressure (CPAP)** - the patient is allowed to breathe spontaneously in the presence of constant (low) airway pressure. This mode may be used to keep the airway open (in obstructive lung disease), collapsed parts of the lung inflated, and/or to help reduce lung fluid. The continuous pressure is often used together with other ventilator modes (such as pressure control and pressure support) to avoid repeated opening and closing of lung alveoli, a potential cause of ventilator-associated lung injury.

Table 2.1: Mechanical ventilator modes

Mode	Description	V/P control	Level of support
CMV	Continuous Mandatory Ventilation	Volume	Controls breathing
A/C	Assist Control Ventilation	Volume	Assists breathing
IMV	Intermittent Mandatory Ventilation	Volume	Spontaneous breathing
PC	Pressure Control	Pressure	Controls breathing
PS	Pressure Support	Pressure	Assists breathing
CPAP	Continuous Positive Airway Pressure	Pressure	Spontaneous breathing

2.2.2 Mechanical ventilator settings

In addition to ventilator mode, there are a variety of settings used to customize ventilator treatment. Each variable is set or observed depending on the ventilator mode and breathing status of the patient. These settings are described here and summarized in Table 2.2.

- **Tidal volume** (V_T) - the volume of air delivered during one inspiration. Historically, a V_T of 10 to 15 mL/kg of predicted body weight has been used. However, lower volumes 6 mL/kg are now recommended for patients with ARDS or pulmonary edema because these lungs have lower respiratory compliance. In volume-control modes, V_T is an adjustable variable. In pressure-control modes, it is observed and is a function of ventilator pressures and a patient's lung compliance.

- **Peak inspiratory pressure** (PIP) - the maximum applied pressure at the airway during inspiration. PIP is the sum of three pressures: the positive end-expiratory pressure ($PEEP$), the pressure due to lung inflation (elastic pressure), and pressure needed to overcome airway resistance as shown in Figure 2.3. PIP, also known as *peak pressure*, is set in pressure-control modes and observed in volume-control modes.

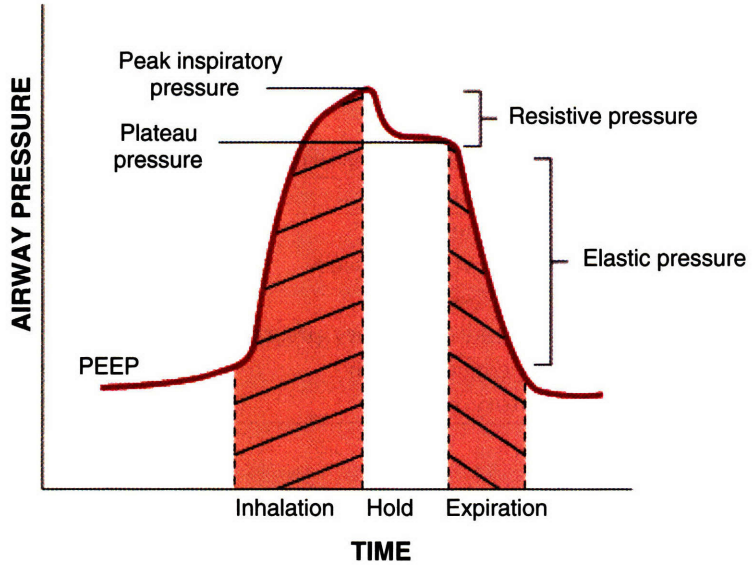


Figure 2.3: Airway pressures during mechanical ventilation. Ventilator settings shown include peak inspiratory pressure, plateau pressure, and PEEP.

- **Plateau pressure** (P_{plat}) - the airway pressure measured immediately after the end of inspiration and before expiration, a period known as the *inspiratory pause* or *inspiratory hold*. P_{plat} is the most direct measurement of the pressures sustained by lung alveoli (because it is recorded when net airflow is zero), so it has been recommended to keep P_{plat} below a certain threshold (28 cmH_2O) to avoid barotrauma. P_{plat} is well correlated with PIP in most patients where airway resistance remains fairly constant.

- **Positive end-expiratory pressure** ($PEEP$) - the airway pressure measured at the end of expiration. A small amount of $PEEP$ (5 to 10 cmH_2O) is recommended to minimize injury associated with repeated opening and closing of lung alveoli. Higher $PEEP$ is sometimes used to recruit collapsed areas of the

lung. This variable can be set in both volume-control and pressure-control modes.

- **Oxygen fraction (FiO_2)** - the fraction of oxygen in inspired air, represented as a number between 0 and 1. Normal air contains 21% oxygen, giving $FiO_2 = 0.21$. Mechanical ventilators can deliver oxygen levels of up to 100%, although FiO_2 is usually kept lower to reduce the chance of oxygen toxicity.

- **Respiratory rate (RR)** - the frequency of breaths delivered by the ventilator. Respiratory rate in normal adults ranges between 10 and 20 breaths per minute. In mechanically ventilated patients, a RR of 8 - 12 breaths per minute is recommended for those without metabolic acidosis. Higher respiratory rates allow less time for exhalation, a problem for patients with obstructive airway disease. Respiratory rate and tidal volume can be adjusted to control minute volume, the volume air delivered per minute:

$$\text{Minute volume } (\frac{\text{mL}}{\text{min}}) = RR * V_T$$

Table 2.2: Mechanical ventilator settings

Setting	Description	Units
V_T	Tidal volume	mL
PIP	Peak inspiratory pressure	cmH_2O
P_{plat}	Plateau pressure	cmH_2O
$PEEP$	Positive end-expiratory pressure	cmH_2O
FiO_2	Oxygen concentration	fraction
RR	Respiratory rate	breaths/min

2.2.3 Advantages of mechanical ventilation

The advantages of mechanical ventilation lie in its ability to provide life-saving therapies in the short term. First, it can be used to inflate parts of the lungs that are collapsed due to atelectasis and provide necessary aeration to the bilateral lungs. Second, the ability to control the ventilation and amount of oxygen delivered makes it possible to correct life-threatening hypoxia and hypercapnia. Third, positive pressures may be used to push out fluid that accumulates in the alveoli, for example to decrease pulmonary edema caused by heart failure. However, this technique does not apply to ARDS, in which lung infiltrates include cells and proteins in addition to fluid from the bloodstream. Finally, intubation for mechanical ventilation may be used to control/protect a patient's airway as a precautionary measure. For example, patients with head injury, in post-operative recovery, or drug overdose may have an impaired respiratory drive and be intubated in anticipation of the need for life support and/or to protect the airway. Thus it is possible for patients without respiratory failure to be intubated for mechanical ventilation.

2.2.4 Disadvantages of mechanical ventilation

Although it is a life-saving technique, mechanical ventilation also has disadvantages because of its invasive nature. First, the use of this therapy prolongs hospital stay because it is necessary to "wean" a patient before disconnecting the ventilator. Such weaning may require several days depending on the patient's ability to recover. Second, there are numerous complications associated with mechanical ventilation, including the risk for pneumothorax (punctured lung), ventilator-associated pneumonia, alveolar injury, and airway injury due to improper intubation or ventilation. The relationship between mechanical ventilation and lung injury (including ARDS) is still under investigation, and this thesis aims to contribute to such an effort.

2.3 Clinical studies on mechanical ventilation and ARDS

2.3.1 The use of mechanical ventilation in ARDS patients

Mechanical ventilation has been an important component of the care of patients with respiratory failure, and it is clear that this therapy was critical to their survival. Traditionally, tidal volumes of 10 to 15 mL/kg predicted body weight (PBW) have been used in patients with respiratory failure [4]. However, it has become apparent that ARDS significantly reduces the amount of normally aerated lung tissues and that high tidal volumes may over-distend the injured lungs [5]. Various clinical trials have thus tried to examine the relationship between ventilator settings and the outcome of ARDS patients (measured as hospital mortality, duration of mechanical ventilation, and duration of non-pulmonary organ failure).

Four randomized controlled trials were conducted in the late 1990s to evaluate the benefit of low vs. traditional tidal volumes in ARDS patients. One of the studies found a significant difference in hospital mortality between patient groups (38% for $V_T \leq 6$ mL/kg PBW vs. 71% for $V_T = 12$ mL/kg PBW, $p = 0.001$) [6]. The other three trials did not find significant differences in patient mortality, possibly because the difference between tidal volumes was not as large ($V_T \leq 8$ mL/kg PBW in low tidal volume groups) [7, 8, 9]. All of these studies had low statistical power due to small sample sizes ($n = 52$ to 120), so a large prospective trial was conducted over three years to address the conflicting results. This trial enrolled 861 patients in 10 institutions and found that lower tidal volumes ($V_T \leq 6$ mL/kg PBW vs. $V_T > 12$ mL/kg PBW) were significantly associated with lower hospital mortality (31% vs. 39.8%, $p = 0.007$) [10].

There has also been discussion of the protective nature of PEEP in patients with respiratory failure. It is known that repeated opening and closing of alveoli during respiratory cycle can promote lung injury in animal models [11, 12]. Thus it has been proposed that PEEP may be used to prevent compression

atelectasis and limit phasic collapse of airways. However, a large randomized trial to examine the effects of high vs. low PEEP (15 cmH_2O vs. 9 cmH_2O) in ARDS patients produced no difference in mortality, duration of mechanical ventilation, or duration of non-pulmonary organ failure [13]. Although PEEP has been used to reduce lung fluid in mild cardiogenic pulmonary edema, it does not reduce lung infiltrates in ARDS [14, 15]. Thus, low tidal volume is the only method of mechanical ventilation that, to date, has been shown to improve survival in patients with ARDS in randomized controlled trials.

2.3.2 Ventilator-induced ARDS

There is strong evidence that mechanical ventilation with high tidal volumes and airway pressures can trigger inflammatory pulmonary edema in animal models [16, 17, 18, 19]. This causes concern for treatment of human patients who are mechanically ventilated but who do not have lung injury at the outset. In fact, patients without respiratory failure make up a significant portion (20 - 30%) of all who are mechanically ventilated in the intensive care unit [20, 21].

Despite the numerous studies on ARDS mortality in humans and lung injury in animals, the evidence for ventilator-induced ARDS in humans is still scarce. It is known that short-term endotracheal intubation and long term mechanical ventilation may increase the risk for nosocomial pneumonia [22]. However, there have been no randomized trials to assess the effects of ventilator settings on new lung injury. The only studies on this topic have been retrospective analyses, which find high tidal volumes to be a risk factor for ALI and ARDS [23, 24]. An important limitation to these studies was that high settings may have been used to correct underlying hypoxia and thus may be an indication of sicker patients. In addition, the relative importance of high airway pressures and high tidal volume as risk factors has not been examined previously in detail.

In general, there have been few studies and no firm recommendations on the optimal settings for patients who require mechanical ventilation for reasons apart from respiratory failure. This thesis aims to investigate this issue through a retrospective analysis of data collected from intensive care units at a single institution hospital. The results of this study contribute to the understanding of ventilator-associated ARDS and has important application in clinical practice.

Chapter 3

Data Extraction and Statistical Methods

This chapter describes the acquisition and analysis of data from the MIMIC-II database. It includes an overview of the database (how and from where the data were collected, types of data available), the algorithms used for patient selection, and the basic theory of odds ratios and logistic regression.

3.1 The MIMIC-II database

The Multi Parameter Intelligent Monitoring of Intensive Care database (MIMIC-II) is a large ICU database developed to support research in intelligent patient monitoring and clinical decision making [25]. It has collected data from intensive care units at Beth Israel Deaconess Medical Center (BIDMC) since 2001, and data acquisition remains an ongoing effort. At the time of research performed for this thesis, MIMIC-II contained over 17,000 electronic medical records for patients admitted between 2001 and 2005.

The MIMIC-II database contains a variety of information from bedside monitors, mechanical ventilators, laboratory tests, progress notes, and recorded medical interventions. Continuous waveform data (ECG, blood pressures, and respiratory waveforms) were obtained from bedside monitors, and vital signs (heart rate, blood pressures, etc) were recorded by ICU nurses on an hourly basis. Ventilator settings were documented by respiratory therapists at the time of intubation and as ventilator settings were adjusted. Blood gas measurements, lab results, IV medications, and fluid I/O were recorded as the interventions were performed. Nursing progress notes were recorded at various times during the patient's hospital stay. Radiological films were evaluated by specialists at the time of patient care, and written evaluations were entered into the database along with the report type and dates. ICD-9 codes were recorded for specific diseases as required by hospital staff upon patient discharge.

3.2 Finding patients of interest

To study the effects of mechanical ventilation on the development of ARDS, we looked for patients from the database who were on the ventilator for longer than 48 hours and who did not have ARDS at the outset. To rule out cardiogenic causes of pulmonary edema, we excluded patients with evidence of congestive heart failure from this study. The remaining patients were then grouped according to the quality of lung health at the beginning of mechanical ventilation. Subsequent development of ARDS was detected by a deterioration of gas exchange and the presence of bilateral infiltrates in the chest x-ray reports.

3.2.1 Calculating the length of mechanical ventilation

The length of mechanical ventilation was defined as the duration of the first continuous ventilation period according to recorded ventilator settings. The most commonly recorded setting was ventilator mode, which was present whenever other ventilator settings were recorded (i.e. tidal volume, respiratory rate). This information was available approximately once every 3 to 10 hours, thus we assumed that ventilator therapy has terminated if 24 hours have passed without a recorded ventilator mode. An algorithm was designed to find the beginning and end points of mechanical ventilation based on this criteria, and an example of this calculation is shown in Figure 3.1. Only patients who were continuously ventilated for greater than 48 hours were included in this study.

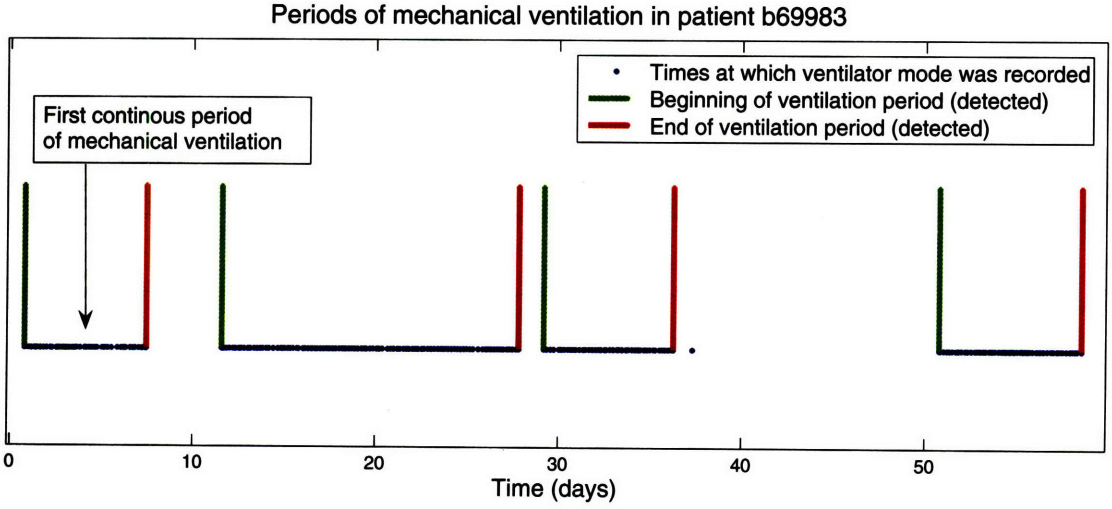


Figure 3.1: Determining the duration of the first continuous mechanical ventilation period

3.2.2 Identifying patients without congestive heart failure

Diagnosis of congestive heart failure (CHF) usually includes a pulmonary artery wedge pressure (PAWP) of greater than 18 mmHg. However, the majority of patients examined in this thesis did not have a recorded PAWP (86%), so patients with CHF were identified using ICD-9 code 428 and were subsequently excluded from the study. Although the accuracy of using ICD-9 codes to identify CHF has not been properly tested, this method has generally been accepted for retrospective clinical studies [26, 27, 28].

3.2.3 Calculating the PaO_2/FiO_2 ratio

The PaO_2/FiO_2 trend was used to determine the quality of gas exchange in the lungs as a function of time. This trend was calculated by finding the ratio of each PaO_2 blood gas measurement to the nearest FiO_2 before the corresponding blood gas value. An example of PaO_2, FiO_2, and the calculated PaO_2/FiO_2 trend is shown in Figure 3.2. This patient developed hypoxemia refractory to oxygen on the 4th day of mechanical ventilation.

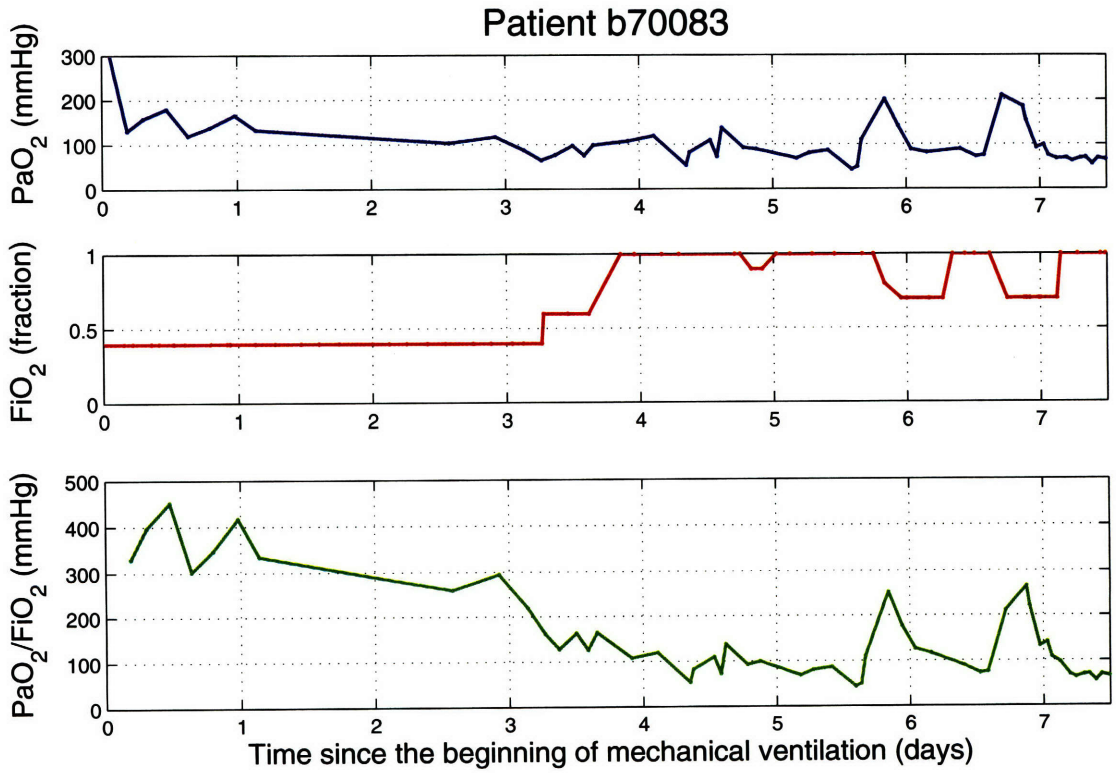


Figure 3.2: Example of the PaO_2/FiO_2 ratio (bottom) calculated from PaO_2 (top) and FiO_2 (middle).

3.2.4 Determining the state of lung health at the beginning of mechanical ventilation

Patients who were mechanically ventilated for more than 48 hours and who did not have evidence of CHF were then categorized into three groups based on their initial lung health: 1) no lung injury, 2) moderate lung injury, or 3) severe lung injury. The PaO_2/FiO_2 criteria for each category is shown in Table 3.1, and patients with severe lung injury at the outset of mechanical ventilation were excluded.

Table 3.1: Categories for initial lung health

Category	Criteria during first 12hrs of mechanical ventilation
Healthy lungs	2 or more $PaO_2/FiO_2 > 300$ mmHg, and 1 or less $PaO_2/FiO_2 < 300$ mmHg
Moderate lung injury (ALI)	2 or more $PaO_2/FiO_2 < 300$ mmHg, and 2 or more $PaO_2/FiO_2 > 200$ mmHg, and 1 or less $PaO_2/FiO_2 < 200$ mmHg
Severe lung injury (ARDS)	2 or more $PaO_2/FiO_2 < 200$ mmHg

3.3 Detecting the onset of ALI and ARDS

To study the risk factors for ALI and ARDS, we examined the following two outcomes: A) the development of acute lung injury (ALI) in patients with healthy lungs at the outset, and B) the development of acute respiratory distress syndrome (ARDS) in patients with healthy or moderately injured lungs at the outset. We have excluded congestive heart failure as a potential cause for hypoxemia, so it remains for us to detect an acute drop in PaO_2/FiO_2 ratio and the appearance of bilateral infiltrates in chest x-rays. These outcomes and their corresponding criteria are summarized in Table 3.2.

Table 3.2: Outcomes of interest

Outcome	Initial lung health	Criteria
ALI	healthy lungs	Not congestive heart failure based on ICD-9 codes PaO_2/FiO_2 drops < 300 mmHg for 24hrs Bilateral infiltrates/consolidations in chest x-ray reports
ARDS	healthy lungs or moderate lung injury	Not congestive heart failure based on ICD-9 codes PaO_2/FiO_2 drops < 200 mmHg for 24hrs Bilateral infiltrates/consolidations in chest x-ray reports

3.3.1 Detecting a deterioration in gas exchange

In patients without lung injury at the outset, we look for the development ALI defined by a drop of PaO_2/FiO_2 below 300 mmHg for 24 hours (note: this also includes the outcome that patients with healthy lungs later develop ARDS). In patients with healthy or moderately injured lungs, we look for the development of ARDS as characterized by a drop in PaO_2/FiO_2 below 200 mmHg for at least 24 hours. The two distinctions are made to independently asses the progression of healthy to injured lungs, and healthy or partially injured to severely injured lungs. An example of PaO_2/FiO_2 trend that represents the development of ALI is shown in Figure 3.3.

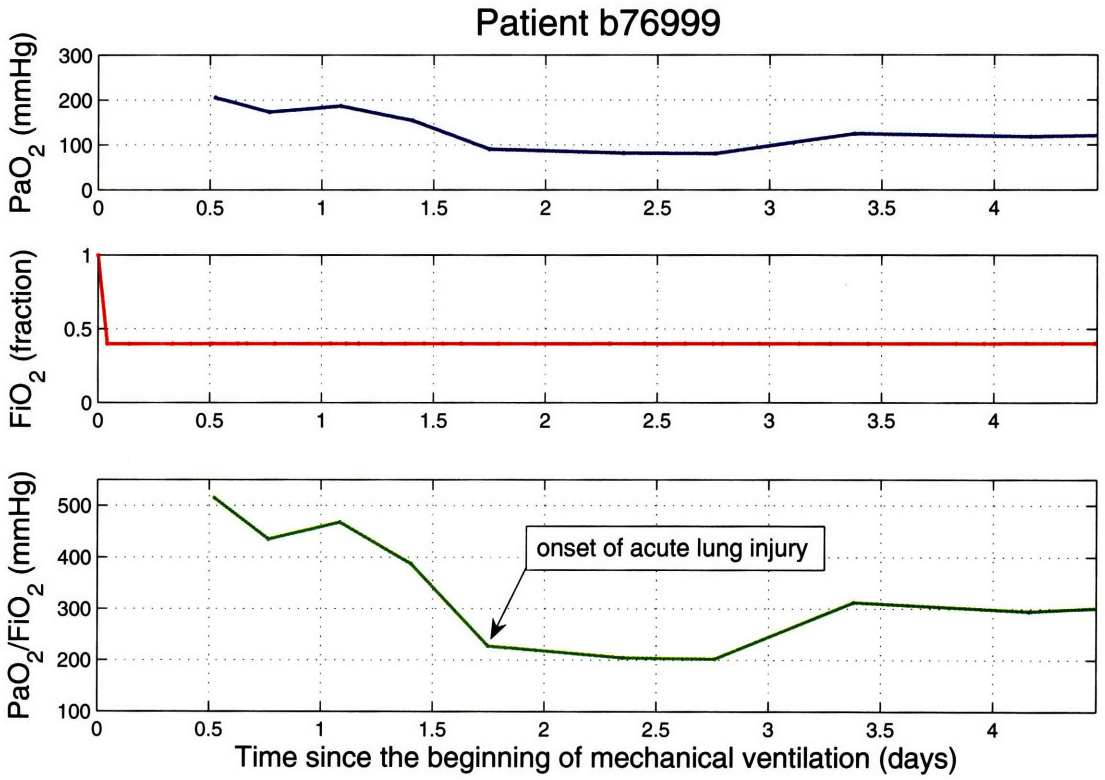


Figure 3.3: An example of gas exchange deterioration in ALI, indicated by a drop in PaO_2/FiO_2 below 300 mmHg for 24hrs or more.

3.3.2 Finding bilateral infiltrates in chest x-ray reports

In patients with a deteriorating gas exchange, chest x-ray (text) reports from 24 hrs before to 72 hours after the drop in PaO_2/FiO_2 ratio were assessed for the presence of bilateral infiltrates and/or lung consolidations. A patient had infiltrates if the report described "opacities," "haziness," "edema," "inflamed,"

"densities," "ARDS," etc. They had lung consolidations if "consolidation," "atelectasis," or "collapse" was present. Infiltrates in one lung and atelectasis in the other was considered consistent with ALI/ARDS. On average, approximately four reports were evaluated for each patient who had a deterioration in gas exchange.

To aid in the diagnosis of bilateral disease, an algorithm was developed to automatically extract information from the chest x-ray reports. This algorithm looked for bilateral infiltrates/consolidations by searching for specific phrases the same way a human reader would. The diagnoses made by this algorithm were verified manually by two expert intensivists (and a graduate student), and discrepancies were settled in a joint reading of the reports. The manual evaluations were later used as a gold standard to assess the performance of the text analysis algorithm; a detailed description of the design and evaluation of this algorithm is presented in Appendix A. Patients were diagnosed with ALI or ARDS only if the drop in PaO_2/FiO_2 had a corresponding chest x-ray report that indicated the presence of bilateral infiltrates/consolidations.

3.4 Extracting data for analysis

3.4.1 Data variables

Once the patient cohort was identified, physiologic information and ventilator settings were collected from the first 24 hours of mechanical ventilation for all patients who were on the ventilator for $\geq$ 48 hrs, who did not have CHF, and who did not have ARDS at the onset of mechanical ventilation. If new lung injury occurred on the first day, data were collected prior to the new injury. The potential risk factors for ALI and ARDS included demographic variables (sex, age, weight, height), indicators of organ health and underlying illness (SAP score, creatinine, ALT, pneumonia, sepsis), ventilator settings (V_T, P_{Plat}, PIP, $PEEP$, FiO_2, RR) and indicators of gas exchange and metabolism (arterial pH, PaO_2, $PaCO_2$, bicarbonate and lactate). Table 3.3 lists all the variables extracted for statistical analyses. When more than one ventilator setting was present on a given day, the "worst" values (highest tidal volume, highest ventilator pressures) were selected. For all non-ventilator variables, the first value after the outset of ventilation was collected. Presence of pneumonia and sepsis as an underlying illness was identified by ICD-9 codes (480 - 486 for pneumonia, 038 for sepsis).

3.4.2 Calculated variables

Calculated variables include the predicted body weight (PBW), normalized tidal volume, static respiratory compliance, SAP score, and PaO_2/FiO_2 ratio. When patient height information was available (60% of

Table 3.3: Variables extracted for statistical analysis

Variable	Units	Description
Demographics		
sex	M = 1, F = 0	patient sex
age	years	patient age
weight	kilograms	patient weight
height	inches	patient height
Organ health		
SAPS	SAP units	simplified acute physiology score
creatinine	mg/dL	a measure of kidney function
ALT	units/L	alanine aminotransferase, a liver enzyme
pneumonia	present=1, absent=0	pulmonary edema due to a lung infection
sepsis	present=1, absent=0	severe immune response to an infection
C_{rs}	mL/cmH_2O	static respiratory compliance
Ventilator settings		
V_T	mL	set tidal volume
V_T/PBW	mL/kg	tidal volume per predicted body weight
P_{Plat}	cmH_2O	plateau pressure
PIP	cmH_2O	peak inspiratory pressure
$PEEP$	cmH_2O	positive end-expiratory pressure
FiO_2	fraction	oxygen fraction
RR	breaths/min	total respiratory rate
Gas exchange		
PaO_2	mmHg	partial pressure of O_2 in arterial blood
$PaCO_2$	mmHg	partial pressure of CO_2 in arterial blood
pH	pH	pH of arterial blood
bicarbonate	mmol/L	concentration of HCO_3 in arterial blood
lactate	mmol/L	concentration of lactate in arterial blood

records), predicted body weight was calculated from patient height using the following formulae [23]:

$$PBW_M \text{ in kg} = 50 + 0.91 * (\text{height in cm} - 152.4) \tag{3.1}$$

$$PBW_F \text{ in kg} = 45.5 + 0.91 * (\text{height in cm} - 152.4) \tag{3.2}$$

The normalized tidal volume was then calculated using predicted body weight:

$$\text{Normalized tidal volume} = \frac{V_T}{PBW} \tag{3.3}$$

The static respiratory compliance C_{rs} (a measure of lung elasticity) was calculated as:

$$C_{rs} = \frac{V_T}{P_{plat} - PEEP} \tag{3.4}$$

The Simplified Acute Physiology Score (SAPS) was determined using physiologic data from the first 24 hours of admission. This data included age, heart rate, systolic ABP, temperature, respiratory rate, urine output, BUN, hematocrit, white blood count, glucose, potassium, sodium, bicarbonate, and Glasgow Coma score. Each variable is mapped to a number between 0 and 4, and the scores are summed to give the SAP score. The guidelines used to calculate SAPS is shown in Table 3.4.

Table 3.4: Variables and values used to calculate SAPS I score.

Variable	**add 0**	**add 1**	**add 2**	**add 3**	**add 4**
Age (years)	≤ 45	45 - 55	55 - 65	65 - 75	> 75
HR (beats/min)	70 - 110		55 - 70 110 - 140	40 - 55 140 - 180	≤ 40 > 180
ABP sys (mmHg)	80 - 150		55 - 80 150 - 190		< 55 > 190
Temp (°C)	36 - 38.4	34 - 36 38.4 - 38.9	32 - 34	30 - 32 38.9 - 41	< 30 > 41
Resp rate (breaths/min)	11 - 24	9 - 11 24 - 34	6 - 9 cpap or vent	34 - 49	< 6 > 50
Urine output (L/day)	0.7 - 3.5	3.5 - 5	0.5 - 0.6 > 5	0.2 - 0.5	< 0.2
BUN (mg/dL)	10 - 20	< 10 20 - 80	80 - 100	100 - 155	> 155
Hematocrit (%)	30 - 46	46 - 50	20 - 30 50 - 60		< 20 > 60
WBC count ($10^3/mm^3$)	3 - 15	15 - 20	1 - 3 20 - 40		< 1 > 40
Glucose (mg/dL)	70 - 250	250 - 500	50 - 70	30 - 50 500 - 800	< 30 > 800
Potassium (mEq/L)	3.5 - 5.5	3 - 3.5 5.5 - 6	2.5 - 3	6 - 7	< 2.5 > 7
Sodium (mEq/L)	130 - 150	150 - 155	120 - 130 155 - 160	110 - 120 160 - 180	< 110 > 180
HCO_3 (mEq/L)	20 - 30	10 - 20 30 - 40		10 - 20 > 40	< 5
Glasgow Coma Score	≥ 13	10 - 13	7 - 10	4 - 7	< 4

3.5 Statistical analyses - odds ratios and logistical regression

This section explains the theory behind odds ratios, logistic regressions, and the p-value. The odds ratio is a useful tool for examining the relationship between a variable and a binary ("yes or no") outcome, and at the same time providing a confidence interval for the significance of that relationship [29]. The variable being examined can be either a continuous variable (i.e. age) or binary variable (i.e. sex), which give us tremendous flexibility in the type of data analysis performed. More importantly, odds ratios in logistical regression allow us to examine the relative importance of variables in affecting an outcome. For example, logistic regressions ultimately help us answer the question, "is tidal volume or airway pressure more important in the development of ventilator-associated lung injury?"

3.5.1 The odds ratio

6-sided die 4-sided die

Figure 3.4: Illustrating the odds ratio using two dice.

The following example illustrates the concept of odds ratio. Consider rolling a normal 6-sided die 60 times to produce, in this case, a total of 10 *ones*. From this observation, the probability of obtaining a *one* is 10/60. However, the odds of rolling a *one* is $\frac{10}{50} = \frac{1}{5}$, because the odds is defined as:

$$Odds = \frac{\#\text{ of successes}}{\#\text{ of failures}} \tag{3.5}$$

Now we roll a 4-sided die 60 times which, by chance, produces 15 *ones*. This result is shown in Table 3.5.

Table 3.5: Hypothetical results from rolling a 6-sided and 4-sided die

	# of *ones*	**# of other outcomes**
6-sided die	10	50
4-sided die	15	45

The odds of rolling a *one* on the 4-sided die is $\frac{15}{45} = \frac{1}{3}$. We can use the odds ratio (OR) to examine the relationship between choosing the 4-sided die and obtaining the result *one*. The "ratio of the odds" is

calculated as the odds for rolling a *one* using the 4-sided die divided by the odds associated with rolling a *one* using the 6-sided die :

$$OR = \frac{\text{odds of rolling } one \text{ using 4-sided die}}{\text{odds of rolling } one \text{ using 6-sided die}} = \frac{15/45}{10/50} = \frac{5}{3} = 1.\overline{6} \tag{3.6}$$

This calculation tells us that there is a higher odds associated with throwing a *one* if we pick the 4-sided die as opposed to the 6-sided die. In general, there is a positive association between two observations if the odds ratio is greater than 1 and a negative association when the odds ratio is between 0 and 1. Note that we obtain the same odds ratio when asking the reverse question, "what are the odds that we picked the 4-sided die given that the roll resulted in a *one*?" The odds ratio remains:

$$OR = \frac{15/10}{45/50} = \frac{5}{3} \tag{3.7}$$

This property makes the odds ratio a useful indicator of the strength of relationship between two observations. The general formula to calculate an odds ratio is given as:

$$OR = \frac{\text{odds of "success" in case 1}}{\text{odds of "success" in case 2}} \tag{3.8}$$

3.5.2 Logistic regression

Logistic regression is a method by which we can examine the relationship between predictor variables and a binary ("yes/no") outcome. In the dice example we calculated the odds ratio between two binary events (picking the 4-sided die and rolling a *one*). When the predictor variable is continuous (ie. patient age) rather than binary (choosing 4-sided or 6-sided die), a logistic regression model can be used to calculate the odds ratio associated with a certain change in the variable.

In univariate logistic regression, a single variable X is used to estimate the probability of success p using the following formula:

$$ln(\frac{p}{1-p}) = \alpha + \beta X \tag{3.9}$$

Given some observed data (a set of X and associated outcomes), the optimal values for α and β are calculated to best fit the data. Various statistical software packages are available to determine these coefficients, and Matlab was used for the calculations in this thesis. If X is a binary variable, it can be represented as 0's and 1's in the logistic regression model. When we solve for p in equation 3.9, we obtain an expression for the estimated probability:

$$p = \frac{1}{1 + e^{-(\alpha+\beta X)}} \tag{3.10}$$

The graph of the estimated probability has a sigmoidal shape, as shown in Figure 3.5. A log plot of the odds of success ($\frac{p}{1-p}$) as a function X is also shown in Figure 3.5.

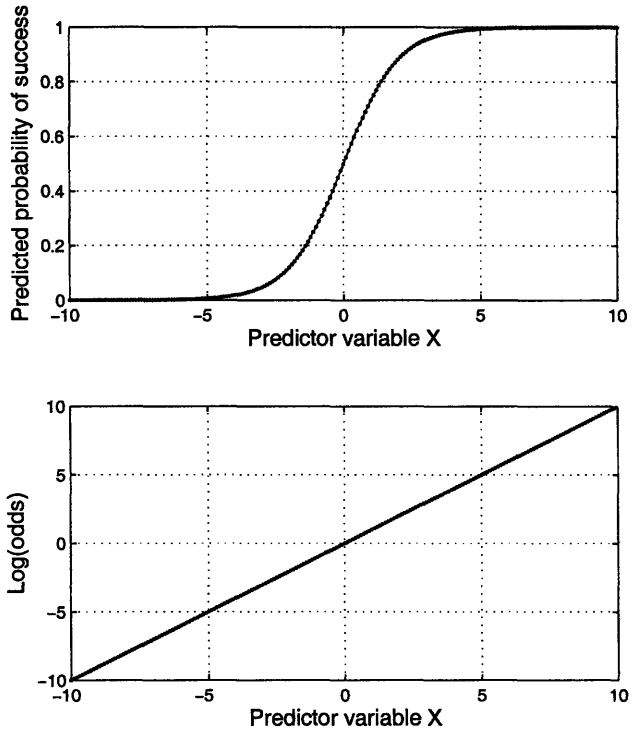


Figure 3.5: Logistic regression curve has a sigmoidal shape, here $\alpha = 0, \beta = 1$, giving $p = \frac{1}{1+e^{-X}}$ (top). The odds of success $\frac{p}{1-p}$ (shown in log scale) increases exponentially as a function of X (bottom).

3.5.3 Odds ratios in logistical regression

Given the coefficients α and β, we can calculate the odds ratio associated with a positive outcome and particular increase in X, which we call Δx:

$$OR_{\Delta x} = \frac{\text{odds of success at } X + \Delta x}{\text{odds of success at } X} = \frac{e^{\alpha+\beta(X+\Delta x)}}{e^{\alpha+\beta X}} = e^{\beta \Delta x} \tag{3.11}$$

Again, if $OR_{\Delta x} > 1$, the positive outcome is more likely to occur at a higher X. The reverse is true if $OR_{\Delta x} < 1$. To find the 95% confidence intervals for each odds ratio, we need the standard error associated with each variable. The standard error, along with p-values and other performance measures, is calculated by statistical packages that perform the logistic regression. The 95% confidence intervals for the odds ratio is then calculated using the regression coefficient β, the standard error SE, and a given change in the variable Δx:

$$OR_{\Delta x}^{\text{upper}} = e^{(\beta+1.96*SE)*\Delta x} \tag{3.12}$$

$$OR_{\Delta x}^{\text{lower}} = e^{(\beta-1.96*SE)*\Delta x} \tag{3.13}$$

3.5.4 Multivariate logistic regression model

To examine the relationship between multiple variables and a particular binary outcome, we can use a multivariate logistic regression model with the following formula:

$$ln(\frac{p}{1-p}) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \tag{3.14}$$

In this equation, X_i is the value of the i^{th} variable, p is the estimated probability of a particular outcome, and the β_i's are coefficients associated with the i^{th} variable. The optimal α and β's are again obtained computationally, and the odds ratios can be calculated in the same manner as in the univariate regression.

3.5.5 The p-value

It is important to discuss the meaning of the p-value: the p-value is the probability that the results observed were due to chance alone. The smaller the p-value, the higher the significance level and stronger the evidence against the null hypothesis (the idea that the results were due to random chance) [30]. A suggested interpretation of this statistic is shown in Figure 3.6. Historically, p = 0.05 has been used as the threshold below which results may be considered significant [31], and results with p < 0.001 provide even stronger evidence for a significant result. When the p-value is below 0.05, the 95% confidence intervals for odds ratios are both above 1 for a positive relationship and below 1 for a negative relationship.

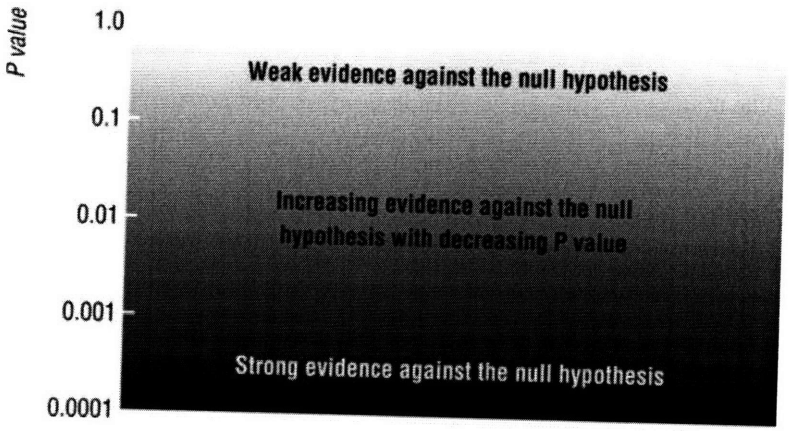


Figure 3.6: Suggested interpretation of p-values from published medical research, adapted from [30].

3.5.6 Performing regressions

Univariate and multivariate logistical regressions were used to find correlations between predictor and outcome variables. The outcomes of interest were development of ALI or ARDS after the onset of mechanical ventilation. In the univariate analysis odds ratios were calculated per one standard deviation increase in predictor variables. Statistically significant variables ($p < 0.05$) were then considered for inclusion in a multivariate model. A forward stepwise and backward stepwise logistic regression were performed to find the optimal model in which all contained variables were statistically significant predictors of ALI and ARDS. SAPS and initial PaO_2/FiO_2 were forced into the multivariate model to control for severity of illness.

To further examine the relative importance of ventilator settings as risk factors for ALI and ARDS, a second multivariable model was created using the following variables: tidal volume, plateau pressure, PEEP, PaO_2/FiO_2, SAPS, and patient weight. In multivariate analyses only, missing values were filled using averages calculated from the entire cohort of patients without ARDS at the onset of mechanical ventilation. All data analyses were performed using Matlab (http://www.mathworks.com/products/matlab/) and its supporting statistical toolboxes.

Chapter 4

Results

This chapter presents the results of statistical analyses for two outcomes of interest: (1) the progression from healthy lungs to acute lung injury (ALI) and (2) the development of acute respiratory distress syndrome (ARDS) in patients with healthy or moderately injured lungs at the outset of mechanical ventilation. Both analyses include a comparison between patients who did and did not develop new ALI/ARDS, univariate and multivariate analyses of potential risk factors, and diagrams to visualize the relationship between risk factors and new lung injury. In addition, the characteristics of the patient cohort are described to give a brief overview of patients who require ventilator therapy in the intensive care unit.

It is important to note that the analyses for the ARDS cohort have more statistical power than that of the ALI cohort simply because the number of patients in the former group is larger. This difference will affect the way results are interpreted, and the discussion of these results are presented in the following chapter.

4.1 Characteristics of the patient cohort

The MIMIC-II database contains a total of 17,493 patients admitted to Beth Israel Deaconess Medical Center (BIDMC) between 2001 and 2005. Of these patients, 2624 required mechanical ventilation for longer than 48 hours, and a subset (1366) did not have congestive heart failure (CHF) during their stay. The average age in these 1366 patients was 59 years and the average length of ICU stay was 13 days. When broken down by location, 29% of patients were in medical ICU, 28% in surgical ICUs, 28% in the cardiac surgery recovery unit, and 15% in coronary care units. The categorization of all patients from the database is shown in Figure 4.1.

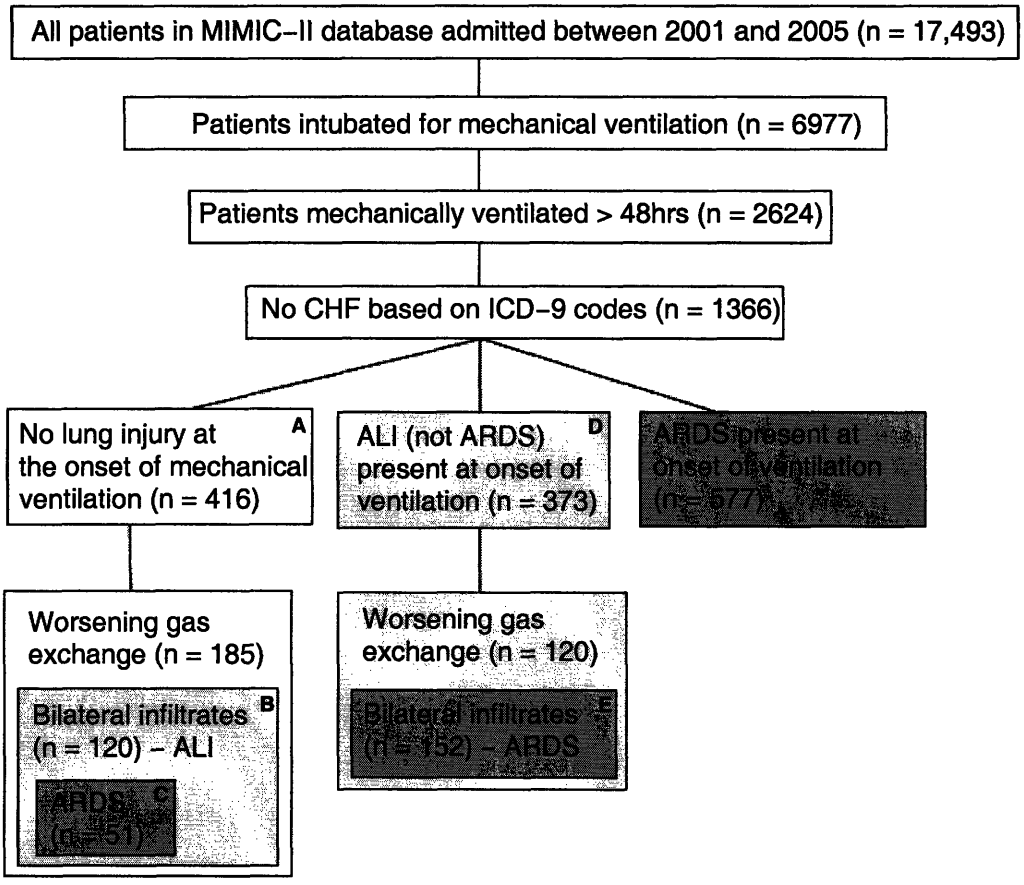


Figure 4.1: Patient distribution from the MIMIC-II database. This study examined the development of ALI in patients initially without lung injury (**A** to **B**), and the development of ARDS in patients initially without ARDS (**A** and **D** to **C** and **E**)

Of the 1366 patients mechanically ventilated for longer than 48hrs and who did not have CHF during their stay, 416 had no lung injury at the onset of mechanical ventilation, 373 had moderate injury, and 577 had severe lung injury. The average length of ventilator therapy increased with the severity of lung sickness (8.6 days in healthy lungs, 9.6 days in moderate injury, 11.2 days for severe injury). Similarly, the average length of ICU stay increased for patients with sicker lungs (11.7 days, 12.3 days, and 14.6 days respectively). The average weight was also higher among patients with sicker lungs (75 kg, 83 kg, and 82 kg respectively). These and other patient characteristics are summarized in Table 4.1.

Table 4.1: Characteristics of patient cohort, grouped by initial lung health.

	All ventilated > 48 hrs (n=1366)[a]	**No lung injury at outset (n=416)[a]**	**ALI at outset (n = 373)[a]**	**ARDS at outset (n = 577)[a]**
Age (years)	59 ± 18 (16 - 99)	58 ± 20 (17 - 99)	62 ± 17 (16 - 91)	58 ± 18 (16 - 92)
Males (n, %)	787, 58%	220, 53%	217, 58%	350, 61%
Height (cm)	170 ± 16.7	168 ± 15	171 ± 22	170 ± 13.3
Weight (kg)	81 ± 23	75 ± 21	83 ± 24	82 ± 22
SAPS[b]	9.3 ± 2.9	9.2 ± 3	9.3 ± 2.8	9.4 ± 3
Days of ICU stay[c]	13.1 ± 10.1	11.7 ± 8.9	12.3 ± 8.7	14.6 ± 11.4
Days of mech. ventilation[d]	10 ± 9.6	8.6 ± 7.6	9.6 ± 8.8	11.2 ± 11

[a]excludes patients with congestive heart failure according to ICD-9 codes; [b]SAPS score based on data from first 24 hours of ICU stay; [c]length of stay in a single care unit; [d]length of first continuous mechanical ventilation period.

4.2 Development of Acute Lung Injury (ALI)

Of 416 patients without lung injury on the first day of mechanical ventilation, 185 had worsening PaO_2/FiO_2, of which 120 (29%) also had bilateral infiltrates and met ALI criteria. The average initial PaO_2/FiO_2 in all patients without lung injury at the outset was 407 ± 101 mmHg, and a histogram of these values is shown in Figure 4.2. The average time until the onset of ALI was 2.6 days from the beginning of mechanical ventilation. Figure 4.3 shows an abnormal case of a patient with healthy lungs at the outset but had one $PaO_2/FiO_2 < 200$ mmHg. Abnormal readings such as this may be due to random error in laboratory tests or from human error.

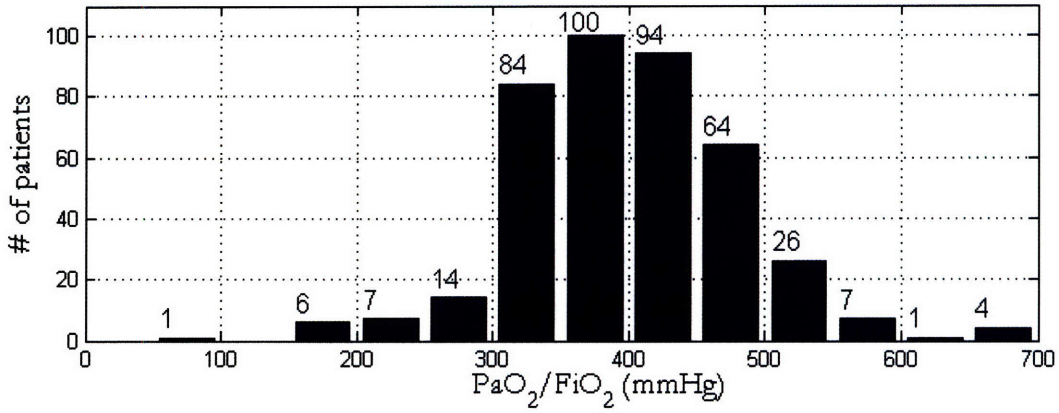


Figure 4.2: Distribution of initial PaO_2/FiO_2 in 416 patients without ALI at the outset of ventilation. The inclusion criteria for this cohort was 2 values $PaO_2/FiO_2 > 300$ and 1 or less $PaO_2/FiO_2 < 300$ mmHg.

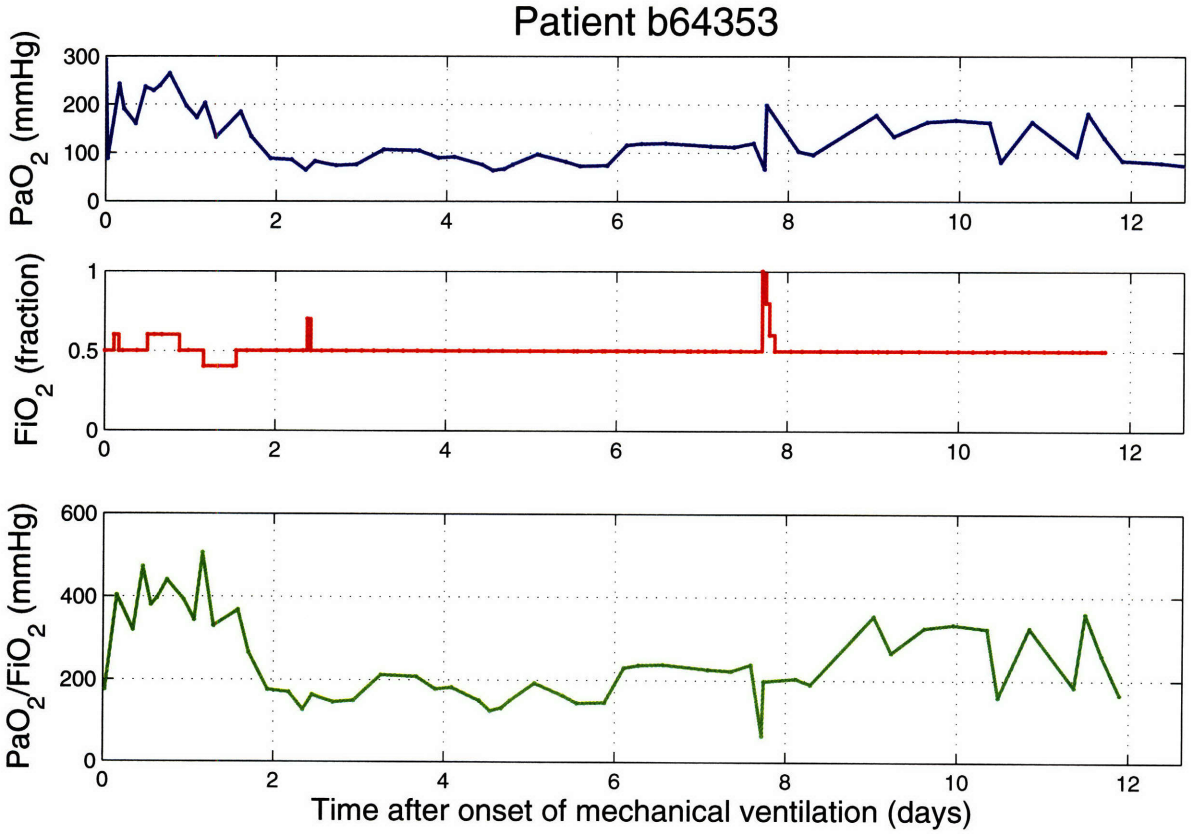


Figure 4.3: Example of patient with initial PaO_2/FiO_2 below 200 mmHg but who did not have lung injury at the outset.

Table 4.2 compares demographic information, severity of illness, and initial ventilator settings between the patients who did and did not eventually develop ALI. On average, patients who acquired ALI required more days of mechanical ventilation (11 days vs. 8 days), had a longer ICU stay (15 days vs. 10 days), had higher weight (80 kg vs. 73 kg), were more likely to have pneumonia (44% vs. 31%), and were more likely to have sepsis (29% vs 15%). The initial PaO_2/FiO_2 ratio was also lower in patients who eventually developed ALI (384 mmHg vs. 417 mmHg).

Table 4.2: Characteristics of 416 patients initially without ALI.

Variable	**Don't develop ALI (n = 296)**	**Develop ALI (n = 120)**
Age (years)	57.9 ± 20.5	58.7 ± 19.9
Sex (male %)	53%	55%
Weight (kg)	73.3 ± 20.2	79.6 ± 22.6
Height (inches)	65.5 ± 6.8	66.9 ± 4.4
Length of stay (days)	**10.5 ± 8**	**14.7 ± 10.2**
Length of ventilation (days)	**7.5 ± 6.6**	**11.2 ± 9.2**
SAPS (SAPS score)	9.1 ± 2.9	9.6 ± 3
V_T (mL)	617 ± 104.6	630.3 ± 116.9
$PEEP$ (cmH_2O)	5.6 ± 1.9	5.9 ± 2.1
PIP (cmH_2O)	28.5 ± 7.7	31.4 ± 9.1
P_{plat} (cmH_2O)	21.2 ± 5.6	23.5 ± 5.8
$\boldsymbol{PaO_2/FiO_2}$ **(mmHg)**	**416.5 ± 107.6**	**383.8 ± 100.8**
$PaCO_2$ (mmHg)	38.3 ± 9	39 ± 11.3
C_{rs} (mL/cmH_2O)	48.5 ± 17.5	44.2 ± 16.5
Pneumonia (%)[a]	**31%**	**44%**
Sepsis (%)[a]	**15%**	**29%**
ARDS (%)[a]	19%	28%

[a]Underlying illnesses according to ICD-9 codes.

4.2.1 Univariate analysis of risk factors for ALI

Univariate continuous-variable logistical regression revealed the following variables to be associated with the development of ALI: $\boldsymbol{P_{plat}}$ (odds ratio 1.5 per standard deviation, 95% confidence interval 1.4 - 1.8), **sepsis** (OR 2.4 for presence of sepsis, 95% CI 1.4 - 3.9), $\boldsymbol{PIP}$ (OR 1.4 per std, 95% CI 1.1 - 1.8), **patient weight** (OR 1.3 per std, 95% CI 1.1 - 1.7), **pneumonia** (OR 1.8 for presence of pneumonia, 95% CI 1.1 - 2.7), **lactate** (OR 1.3 per std, 95% CI 1.1 - 1.6), and $\boldsymbol{C_{rs}}$ (OR 0.76 per std, 95% CI 0.6 - 1.0). V_T and V_T per predicted body weight (p = 0.265 and 0.740) were not found to be significantly associated with new ALI. Table 4.3 summarizes the odds ratios and p-values from univariate logistic regressions.

A visualization of the relationship between day one plateau pressure and new lung injury is shown in Figure 4.4. The percentage of patients who develop ALI increases from 20% to 40% as P_{plat} increases from 16 to

Table 4.3: Univariate analysis of risk factors for ALI in 416 patients without ALI at the outset of mechanical ventilation. Variables are listed in order of statistical significance (lowest to highest p-value).

Variable	Mean	Standard Dev (STD)	Odds Ratio per STD (95% confidence interval)	p-value
P_{plat}	**21.9**	**5.8 cmH_2O**	**1.48 (1.19 - 1.84)**	**< 0.001**
Sepsis[a]	**-**	**-**	**2.36 (1.43 - 3.92)**	**0.001**
PIP	**29.4**	**8.2 cmH_2O**	**1.41 (1.13 - 1.75)**	**0.002**
Weight	**75.2**	**21.1 kg**	**1.34 (1.08 - 1.66)**	**0.008**
Pneumonia[a]	**-**	**-**	**1.76 (1.13 - 2.71)**	**0.012**
Lactate	**2.8**	**2.3 mmol/L**	**1.31 (1.05 - 1.64)**	**0.016**
C_{rs}	**47.3**	**17.3 mL/cmH_2O**	**0.76 (0.60 - 0.97)**	**0.024**
SAPS	9.2	3.0 score	1.20 (0.97 - 1.48)	0.093
Height	66.0	6.1 inches	1.37 (0.94 - 2.00)	0.097
PEEP	5.7	2.0 cmH_2O	1.16 (0.95 - 1.42)	0.151
pH	7.4	0.1 pH	0.86 (0.69 - 1.06)	0.155
Creatinine	1.2	1.2 mg/dL	1.15 (0.94 - 1.40)	0.179
V_T	620.9	108.4 mL	1.13 (0.91 - 1.40)	0.265
Bicarbonate	23.1	5.1 mmol/L	0.91 (0.74 - 1.13)	0.411
ALT	170.1	550.6 units/L	0.89 (0.65 - 1.21)	0.451
$PaCO_2$	38.5	9.7 mmHg	1.08 (0.88 - 1.33)	0.468
Sex (if male)[a]	-	-	0.92 (0.72 - 1.69)	0.640
Age	58.1	20.3 years	1.04 (0.84 - 1.29)	0.733
V_T/PBW	10.2	2.3 mL/kg	0.96 (0.73 - 1.25)	0.740
PaO_2	257.0	115.8 mmHg	1.03 (0.83 - 1.27)	0.804
Resp rate	20.8	6.1 bpm	0.98 (0.79 - 1.21)	0.849

[a] Odds ratio calculated for presence of sepsis, pneumonia, and male sex. SDT, standard deviation; *PIP*, peak inspiratory pressure; P_{plat}, plateau pressure; *PEEP*, positive end-expiratory pressure; V_T, tidal volume.; SAPS, simplified acute physiology score; PBW, predicted body weight; C_{rs}, static respiratory compliance.

32 mmHg. The association between high plateau pressure and incident ALI in the context of other risk factors is explored in multivariate regressions.

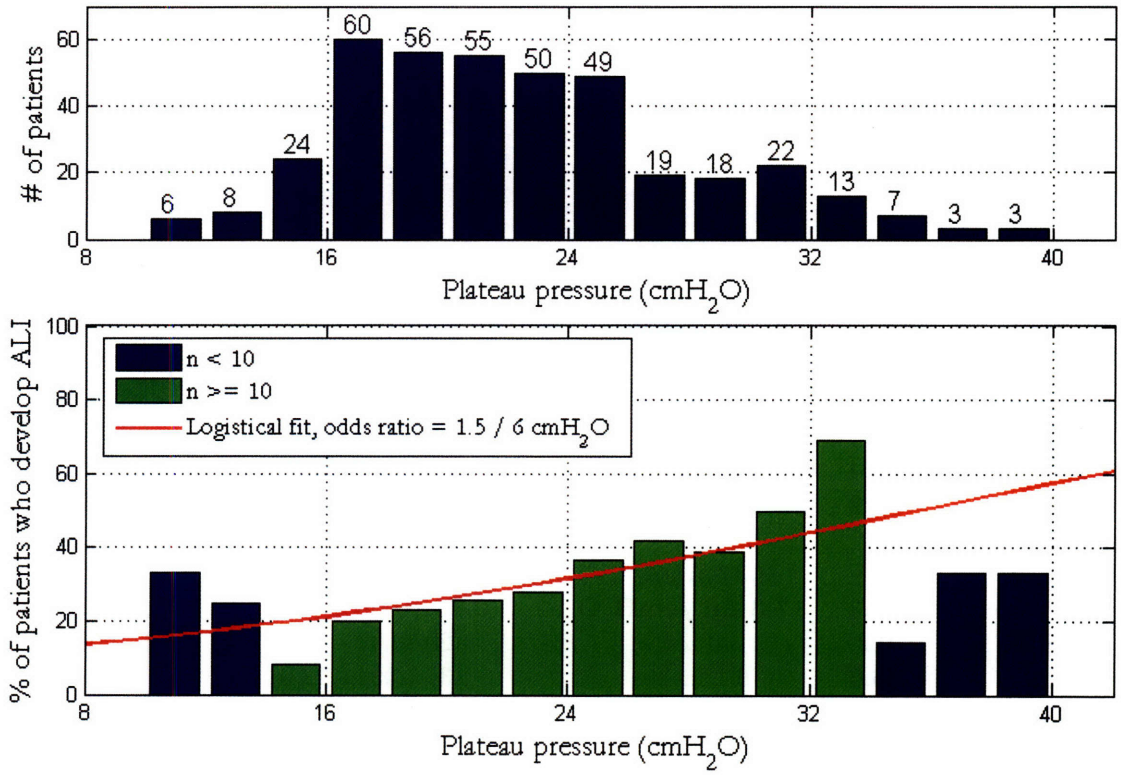


Figure 4.4: Day one plateau pressures in 416 patients without ALI at the onset of mechanical ventilation (top), and the risk of developing ALI as a function of P_{plat} (bottom).

4.2.2 Multivariate analysis of risk factors for ALI

Variables with p-values of less than 0.05 from univariate regressions (P_{plat}, sepsis, PIP, pneumonia, lactate, and C_{rs}) were considered for multivariate regression analysis using a forward-search and backward-search method. Both searches produced an optimal model that included P_{plat}, lactate, and sepsis. PaO_2/FiO_2 and SAPS were then added to model to control for severity of illness, giving the final combination of variables: **sepsis** (OR 1.99, p = 0.011), $\boldsymbol{P_{plat}}$ (OR 1.32 per std, p = 0.014), $\boldsymbol{PaO_2/FiO_2}$ (OR 0.75 per std, p = 0.31), **lactate** (OR 1.23 per std, p = 0.049), and **SAPS** (OR 1.16 per STD, p = 0.185). All variables except for SAPS remained significant predictors of ALI with a p-value < 0.05. This model is shown in Table 4.4.

Another multivariate model was created to examine the relative contributions of the ventilator pressures

and tidal volumes to the development of ALI. This analysis, shown in Table 4.5, included the following variables: V_T, P_{plat}, $PEEP$, patient weight, PaO_2/FiO_2, and SAPS. PIP is known to be correlated with P_{plat} ($R^2 = 0.53$, $p < 0.001$). Including PIP and P_{plat} in the same model decreases the statistical power of both variables as risk factors for lung injury, so only one (P_{plat}) was included. The second multivariate regression shows that $\boldsymbol{P_{plat}}$ (OR 1.34 per std, p = 0.016) and $\boldsymbol{PaO_2/FiO_2}$ (OR 0.75 per std, p = 0.03) remained significant predictors of ALI while V_T, $PEEP$, weight, and SAPS did not ($p > 0.05$).

Table 4.4: Multivariate regression of risk factors for ALI in 416 patients initially without ALI. Optimal model was achieved through forward and backward search of risk factors found in univariate analysis.

Variable	Mean	Standard Dev STD	Odds Ratio per STD (95% confidence interval)	p-value
Sepsis[a]	**-**	**-**	**1.99 (1.16 - 3.36)**	**0.011**
$\boldsymbol{P_{plat}}$	**21.9**	**5.6 cmH_2O**	**1.32 (1.06 - 1.66)**	**0.014**
$\boldsymbol{PaO_2/FiO_2}$	**407.1**	**106.6 mmHg**	**0.75 (0.57 - 0.97)**	**0.031**
Lactate	**2.8**	**2.1 mmol/L**	**1.23 (1.00 - 1.51)**	**0.049**
SAPS	9.2	3.0 SAP score	1.16 (0.93 - 1.45)	0.185

SDT, standard deviation; SAPS, simplified acute physiology score. [a]Odds ratio calculated for presence of sepsis.

Table 4.5: Multivariate analysis of ventilator-associated risk factors for ALI in 416 patients without ALI at the outset. PaO_2/FiO_2 and weight were included to control for severity of illness.

Variable	Mean	Standard Dev	Odds Ratio per STD (95% confidence interval)	p-value
$\boldsymbol{P_{plat}}$	**21.9**	**5.6 cmH_2O**	**1.34 (1.06 - 1.71)**	**0.016**
$\boldsymbol{PaO_2/FiO_2}$	**407.1**	**106.6 mmHg**	**0.75 (0.57 - 0.97)**	**0.030**
SAPS	9.2	3.0 SAP score	1.24 (0.99 - 1.55)	0.056
Weight	75.2	20.5 kg	1.27 (0.98 - 1.64)	0.066
V_T	620.9	106.3 mL	0.95 (0.73 - 1.23)	0.675
$PEEP$	5.7	1.9 cmH_2O	0.99 (0.79 - 1.24)	0.926

SDT, standard deviation; SAPS, simplified acute physiology score.

4.3 Development of Acute Respiratory Distress Syndrome (ARDS)

This section describes the patient cohort used to identify risk factors for ARDS. Of 789 patients with moderate or no lung injury (but not ARDS) at the beginning of mechanical ventilation, 305 had worsening PaO_2/FiO_2, of which 152 (19%) also had bilateral infiltrates and met ARDS criteria. The average PaO_2/FiO_2 at the outset was 302 ± 110 mmHg, and a histogram of these values is shown in Figure 4.5. The average time until the development of ARDS was 3.4 days from the beginning of ventilator therapy.

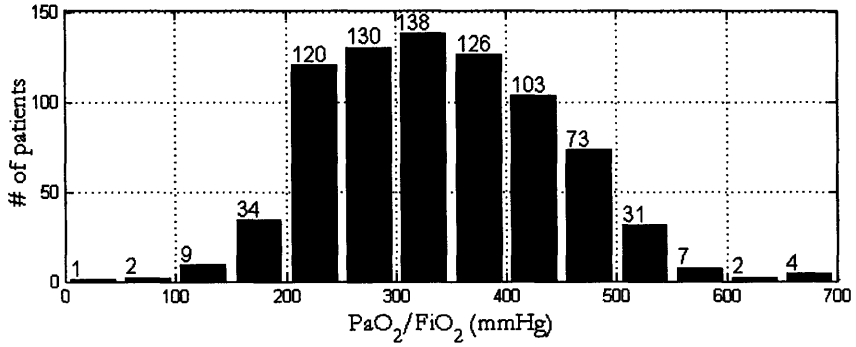


Figure 4.5: Distribution of PaO_2/FiO_2 in 789 patients without ARDS at the outset. Inclusion criteria was 2 $PaO_2/FiO_2 > 200$ and 1 or less $PaO_2/FiO_2 < 200$ mmHg in the first 12hrs of mechanical ventilation.

Table 4.6: Characteristics of 789 patients initially without ARDS.

Variable	Don't develop ARDS (n = 636)	Develop ARDS (n = 152)
Age (years)	60.3 ± 19.1	57.9 ± 18.2
Sex (male %)	54%	50%
Weight (kg)	77.4 ± 22.4	85.7 ± 24.2
Height (inches)	66.4 ± 5.6	67.8 ± 11.8
Length of stay (days)	**11.1 ± 8.1**	**15.5 ± 10.6**
Length of ventilation (days)	**8.2 ± 7.4**	**12.8 ± 10.3**
SAPS (score)	9.2 ± 2.8	9.5 ± 3
V_T (mL)	618 ± 112	651.7 ± 119.1
$PEEP$ (cmH_2O)	5.8 ± 2	6.6 ± 2.8
PIP (cmH_2O)	30 ± 8.2	33.8 ± 9
P_{plat} (cmH_2O)	22.5 ± 6.3	25.4 ± 5.9
PaO_2/FiO_2 (mmHg)	**279 ± 93**	**230.2 ± 85.3**
$PaCO_2$ (mmHg)	39.6 ± 10.9	40.5 ± 11.1
C_{rs} (mL/cmH_2O)	45.7 ± 18.5	41.9 ± 14.7
Pneumonia (%)[a]	**36%**	**44%**
Sepsis (%)[a]	**20%**	**32%**
ARDS (%)[a]	21%	33%

[a]Underlying illnesses according to ICD-9 codes.

Table 4.6 compares demographic information, severity of illness, and day one ventilator settings between patients who did and did not later have acute respiratory distress syndrome. On average, those who developed ARDS required a longer period of mechanical ventilation (12.8 vs. 8.2 days), had a longer hospital stay (15.5 vs. 11.1 days), had a higher incidence of pneumonia (44% vs. 36%) as well as sepsis (32% vs 20%), and had a lower initial PaO_2/FiO_2 (302 vs. 355 mmHg).

4.3.1 Univariate analysis of risk factors for ARDS

Univariate logistic regressions identified the following variables to be well associated with the development of ARDS (Table 4.7): ***PIP*** (OR 1.5 per std, 95% CI 1.3 - 1.8), $\boldsymbol{P_{plat}}$ (OR 1.5 per std, 95% CI 1.3 - 1.9), **patient weight** (OR 1.4 per std, 95% CI 1.2 - 1.7), ***PEEP*** (OR 1.4 per std, 95% CI 1.2 - 1.7), **sepsis** (OR 1.95, 95% CI 1.3 - 2.9), $\boldsymbol{V_T}$ (OR 1.4 per std, 95% CI 1.1 - 1.7), **pneumonia** (OR 1.8, 95% CI 1.2 - 2.7), **blood pH** (OR 0.8 per std, 95% CI 0.6 - 0.9), and **lactate** (OR 1.2 per std, 95%CI 1.0 - 1.4).

Table 4.7: Univariate analysis of risk factors for ARDS in 789 patients without ARDS at the outset of mechanical ventilation. Variables are listed in order of statistical significance (lowest to highest p-value).

Variable	Mean	Standard dev	Odds ratio per STD (95% confidence interval)	p-value
PIP	**30.7**	**8.5 cmH_2O**	**1.53 (1.28 - 1.84)**	**< 0.001**
$\boldsymbol{P_{plat}}$	**23.1**	**6.3 cmH_2O**	**1.54 (1.28 - 1.85)**	**< 0.001**
Weight	**79**	**23.0 kg**	**1.39 (1.18 - 1.65)**	**< 0.001**
PEEP	**5.9**	**2.2 cmH_2O**	**1.35 (1.15 - 1.58)**	**< 0.001**
Sepsis	**-**	**-**	**1.95 (1.31 - 2.88)**	**< 0.001**
$\boldsymbol{V_T}$	**624.6**	**114.2 mL**	**1.36 (1.13 - 1.65)**	**0.001**
Pneumonia	**-**	**-**	**1.82 (1.22 - 2.69)**	**0.002**
pH	**7.4**	**0.1 pH**	**0.77 (0.65 - 0.91)**	**0.003**
Lactate	**2.8**	**2.4 mmol/L**	**1.19 (1.00 - 1.41)**	**0.044**
Creatinine	1.2	1.2 mg/dL	1.14 (0.97 - 1.33)	0.110
Height	66.7	7.6 inches	1.22 (0.95 - 1.56)	0.120
Sex (if male)	-	-	1.32 (0.92 - 1.90)	0.126
Bicarbonate	23.5	5.4 mmol/L	0.88 (0.73 - 1.05)	0.158
Age	59.9	18.9 years	0.88 (0.74 - 1.05)	0.163
PaO_2	118	46.1 mmHg	0.88 (0.73 - 1.06)	0.182
C_{rs}	45.7	39.4 mL/cmH_2O	0.87 (0.70 - 1.08)	0.216
SAPS	9.2	2.9 SAPS	1.11 (0.93 - 1.33)	0.235
$PaCO_2$	39.7	10.9 mmHg	1.09 (0.92 - 1.29)	0.342
V_T/PBW	10.1	2.5 mL/kg	0.94 (0.76 - 1.16)	0.559
ALT	150.1	503.1 units/L	0.95 (0.75 - 1.19)	0.641
Respiratory Rate	21.4	6.2 bpm	0.97 (0.82 - 1.17)	0.779

SDT, standard deviation; PIP, peak inspiratory pressure; P_{plat}, plateau pressure; $PEEP$, positive end-expiratory pressure; V_T, tidal volume.; SAPS, simplified acute physiology score; PBW, predicted body weight.

Variables that were not significant predictors of ARDS included patient age, sex, height, SAPS, V_T/PBW respiratory rate, PaO_2, $PaCO_2$, creatinine, bicarbonate, and C_{rs}. These variables all had a p-value greater than 0.05 in univariate analysis.

It is meaningful to visually examine the relationship between risk factors and subsequent development of lung injury. Figure 4.6 shows the effects of differing day one plateau pressures; the percentage of patients who develop ARDS increases from 10% to 30% as P_{plat} increases from 16 to 32 mmHg. PIP, a ventilator setting known to be well correlated with plateau pressure, exhibits a similar trend as shown in Figure 4.7.

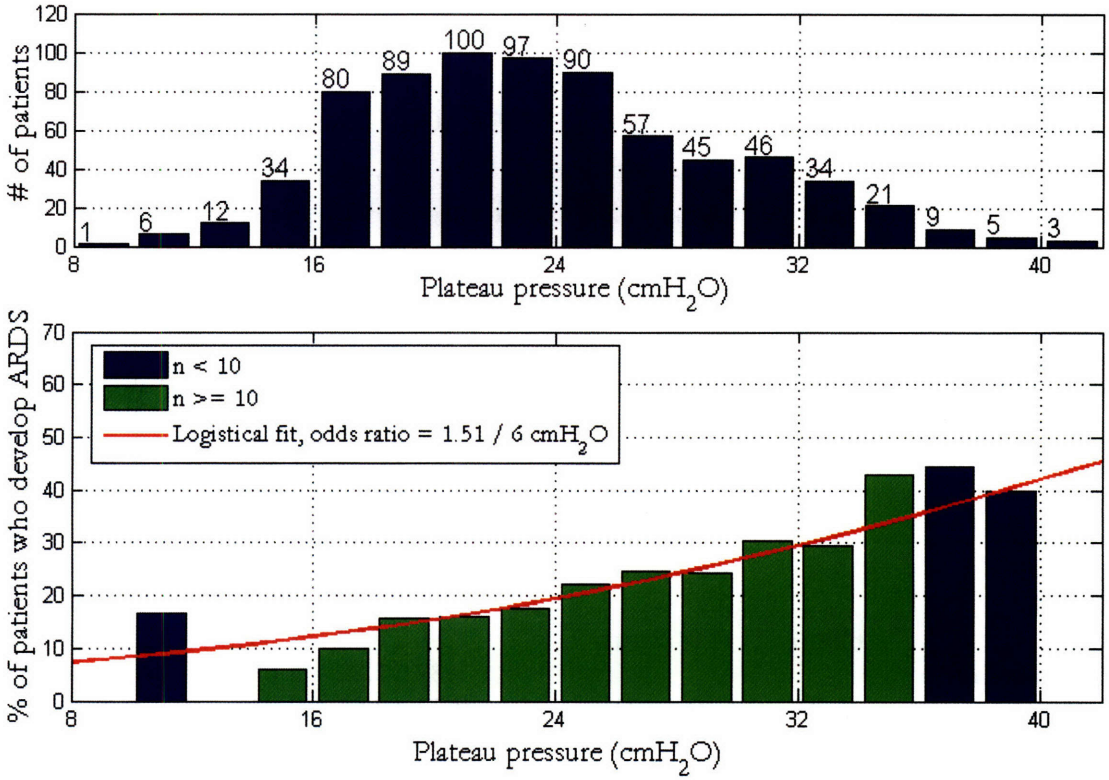


Figure 4.6: Day one plateau pressure in 789 patients without ARDS at the outset of ventilation (top), and the risk of developing ARDS as a function of P_{plat} (bottom).

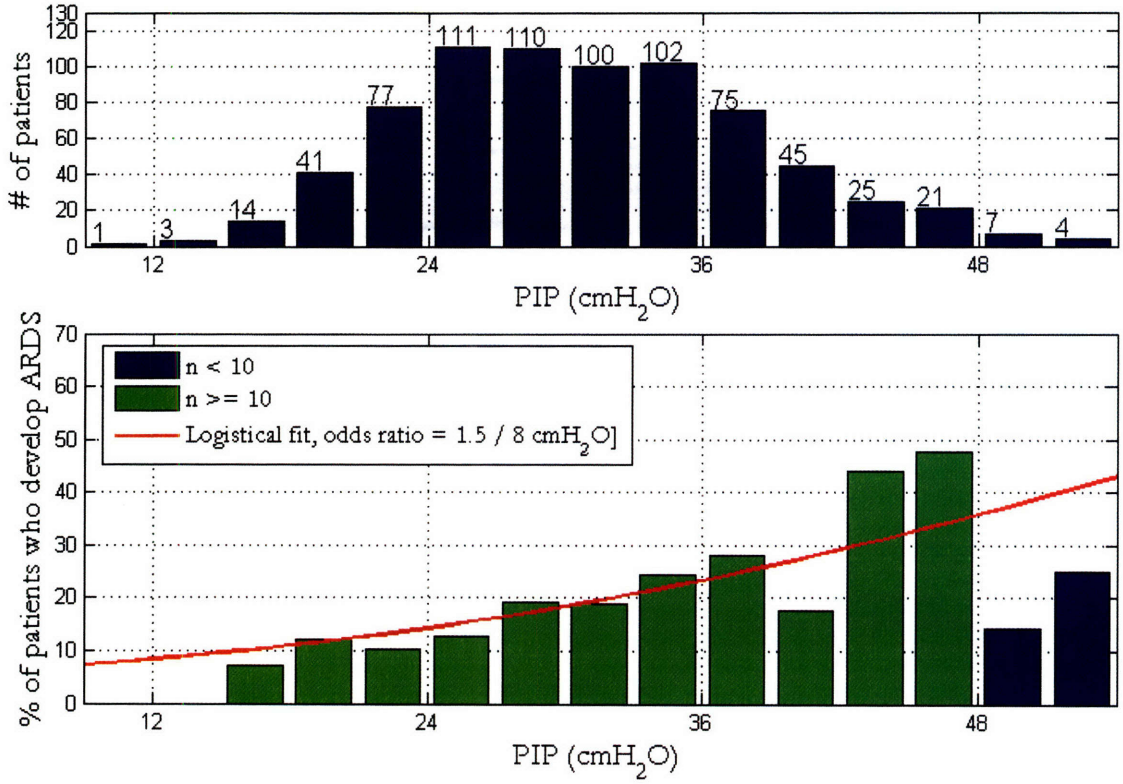


Figure 4.7: Day one peak inspiratory pressure in 789 patients without ARDS at the outset of ventilation (top), and the risk of developing ARDS as a function of PIP (bottom).

There is a more pronounced relationship between initial V_T and development of ARDS (Figure 4.8) than that associated with the development for ALI. Note that the tidal volume is set in intervals of 50 mL, and the percentage of patients who acquire ARDS increases from 10% to 30% as V_T increases from 450 to 800 mL. When examining $PEEP$, we see that most patients are given a $PEEP$ of 5 cmH_2O, and the risk for ARDS increases from 18% to 30% as $PEEP$ increases from 5 to 10 cmH_2O (Figure 4.9).

Two interesting physiologic predictors of ARDS were patient weight and arterial pH. The percent of patients developing ARDS increases from 10% to 30% as patient weight increases from 50 to 100 kg, shown in Figure 4.10. The risk for ARDS increases from 10% to 30% as initial arterial pH decreases from 7.45 to 7.15, however the risk is also high (near 40%) for values of pH above 7.53 (Figure 4.11).

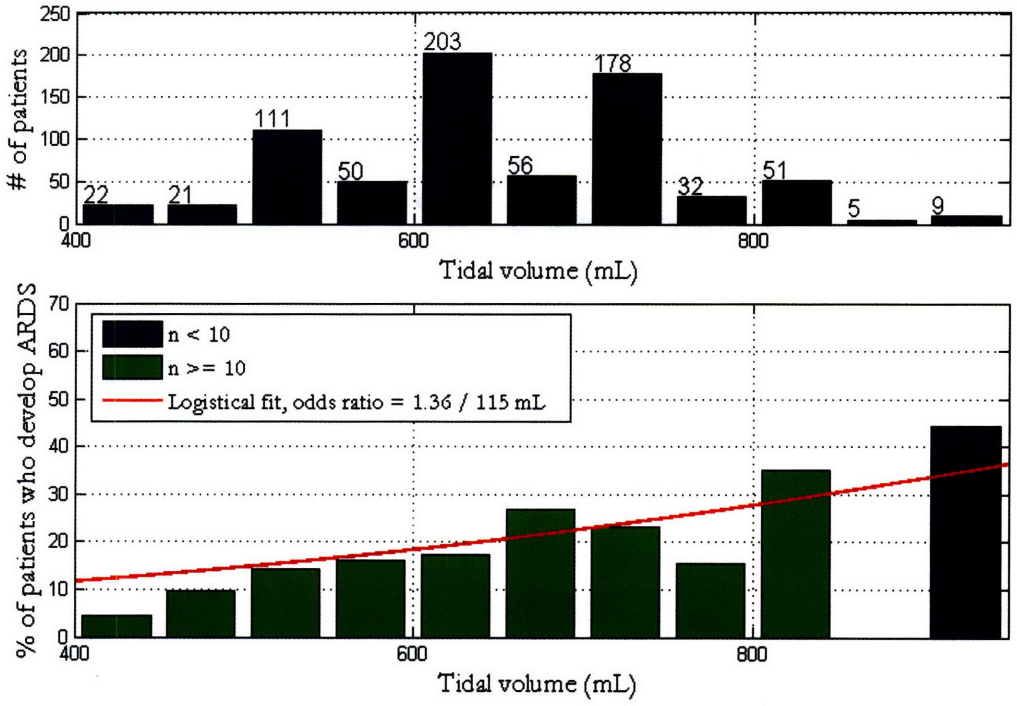


Figure 4.8: Day one tidal volume in 789 patients without ARDS at the outset of ventilation (top), and the risk of developing ARDS as a function of V_T (bottom).

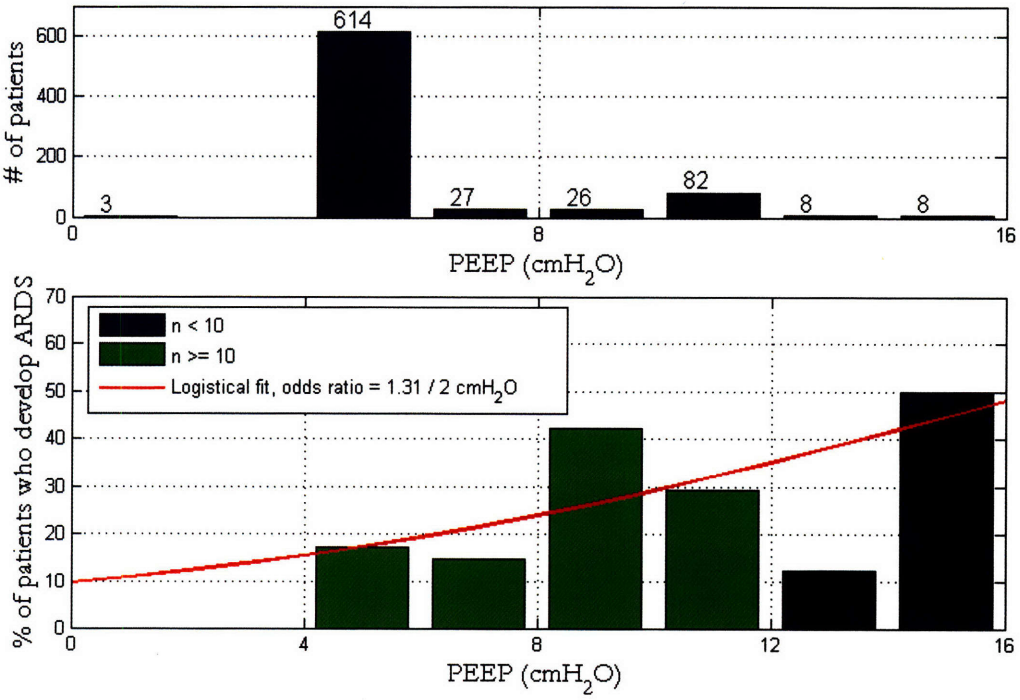


Figure 4.9: Day one positive end-expiratory pressure in 789 patients without ARDS at the outset of ventilation (top), and the risk of developing ARDS as a function of $PEEP$ (bottom).

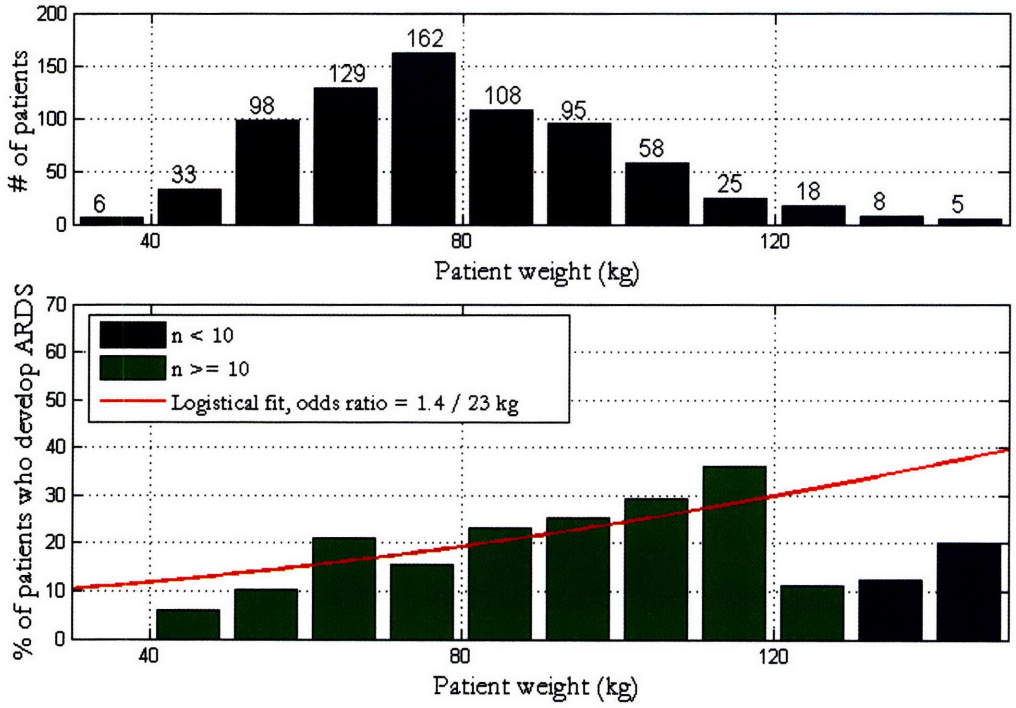


Figure 4.10: Patient weight in 789 patients without ARDS at the outset of ventilation (top), and the risk of developing ARDS as a function of weight (bottom).

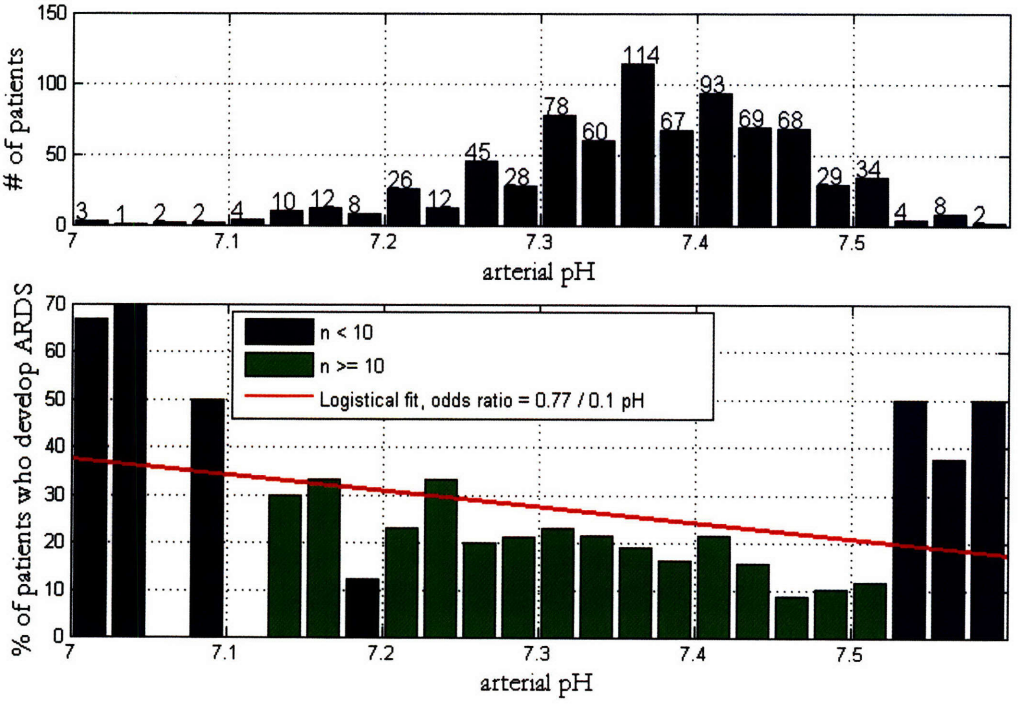


Figure 4.11: Day one arterial pH in 789 patients without ARDS at the outset of ventilation (top), and the risk of developing ARDS as a function of pH (bottom).

4.3.2 Multivariate analysis of risk factors for ARDS

Variables found to be significantly associated with ARDS in univariate analysis (PIP, P_{plat}, weight, $PEEP$, sepsis, V_T, pneumonia, arterial pH, and lactate) were considered for the multivariate regression model. Similar to the analysis for ALI, the optimal multivariate model was found by a forward-search and backward-search on the significant variables. PaO_2/FiO_2 and SAPS were again added to control for severity of illness. Both search methods produced a model that included ***PIP*** (odds ratio 1.35 per standard deviation, p = 0.002), **sepsis** (OR 1.8, p = 0.005), $\boldsymbol{V_T}$ (OR 1.27 per std, p = 0.015), and **pH** (OR 0.83 per std, p = 0.053). This model, shown in Table 4.8, did not include V_T/PBW.

One last multivariate model examined the relative importance of different ventilator settings as risk factors for ARDS (Table 4.9): P_{plat}, $PEEP$, V_T, PIP, patient weight, PaO_2/FiO_2, and SAPS. In this analysis, $\boldsymbol{P_{plat}}$ (OR 1.27 per std, p = 0.018) and $\boldsymbol{PaO_2/FiO_2}$ (OR 0.64 per std, p < 0.001) remained significant predictors while V_T (OR 1.20 per std, p = 0.081), $PEEP$ (OR 1.16, p = 0.083), and weight (OR 1.14, p = 0.195) did not. PIP correlated well with P_{plat} (R^2 = 0.61, p < 0.001) and was not included in this model.

Table 4.8: Multivariate regression of risk factors for ARDS in 789 patients initially without ARDS. Optimal model was achieved through forward and backward search of risk factors found in univariate analysis.

	Mean	Standard Dev	Odds Ratio per STD (95% confidence interval)	p-value
PaO_2/FiO_2	**345.5**	**115.9 mmHg**	**0.62 (0.49 - 0.77)**	**< 0.001**
PIP	**30.7**	**8.2 cmH_2O**	**1.35 (1.12 - 1.62)**	**0.002**
Sepsis	**-**	**-**	**1.81 (1.20 - 2.70)**	**0.005**
V_T	**623.5**	**111.9 mL**	**1.27 (1.05 - 1.55)**	**0.015**
pH	7.4	0.1 pH	0.83 (0.70 - 1.00)	0.053
SAPS	9.2	2.9 score	1.11 (0.92 - 1.33)	0.291

SDT, standard deviation; SAPS, simplified acute physiology score. [a] Odds ratio calculated for presence of underlying illness.

Table 4.9: Multivariate analysis of ventilator-associated risk factors for ARDS in 789 patients without ARDS at the outset. PaO_2/FiO_2 and weight were included to control for severity of illness.

	Mean	Standard Dev	Odds Ratio per STD (95% confidence interval)	p-value
PaO_2/FiO_2	**345.5**	**115.9 mmHg**	**0.64 (0.51 - 0.80)**	**< 0.001**
P_{plat}	**23.1**	**6.1 cmH_2O**	**1.27 (1.04 - 1.54)**	**0.018**
SAPS	9.2	2.9 score	1.18 (0.98 - 1.42)	0.076
V_T	623.5	111.9 mL	1.20 (0.98 - 1.48)	0.081
$PEEP$	5.9	2.2 cmH_2O	1.16 (0.98 - 1.38)	0.083
Weight	79.0	22.4 kg	1.14 (0.94 - 1.38)	0.195

SDT, standard deviation; SAPS, simplified acute physiology score.

Chapter 5

Discussion and Conclusions

5.1 Discussion

This retrospective cohort study sheds new light on the risk factors for ALI and ARDS in patients who were mechanically ventilated for longer than 48 hours in the ICU. In 416 patients who had healthy lungs at the outset, 120 (29%) later developed ALI. Of 789 patients with healthy or moderately injured lungs at the outset, 152 (19%) developed ARDS. Several associations were found between day one ventilator settings and new ALI/ARDS, suggesting that ventilator-associated lung injury may be a preventable illness in some cases. However, the complexity of ARDS and the numerous physiologic risk factors make the disease rather difficult to predict and treat in the ICU. The ventilator settings and physiologic variables associated with new lung injury are discussed in the following sections.

5.1.1 Plateau pressure (P_{plat}) and peak inspiratory pressure (PIP)

Ventilation with high airway pressure is an important risk factor for respiratory failure. In patients without lung injury at the outset, high P_{plat} was significantly associated with the development of ALI in univariate analysis ($p < 0.001$) as well as when adjusted for V_T, $PEEP$, patient weight, PaO_2/FiO_2, sepsis, lactate, and SAPS ($p = 0.016$). Similarly, high P_{plat} (and PIP) was significantly associated with ARDS ($p < 0.001$), even when controlled for V_T, $PEEP$, patient weight, PaO_2/FiO_2, sepsis, SAPS, and pH ($p = 0.018$). P_{plat} is the most direct measurement of the pressures sustained by the lung alveoli. PIP and P_{plat} were well correlated ($R^2 = 0.61$), suggesting that a high PIP is likely to produce an elevated P_{plat} and thus increases the chance for lung overdistention. These results suggest that P_{plat} and PIP are the most critical ventilator-associated risk factors for the development of new ALI and ARDS.

5.1.2 Positive end-expiratory pressure ($PEEP$)

$PEEP$ was not an important risk factor for ALI/ARDS, especially when examined in the presence of P_{plat} and/or PIP. In patients with healthy lungs at the outset, $PEEP$ was not significantly associated with

the development of ALI (p = 0.151). In patients without ARDS at the outset, $PEEP$ was associated new ARDS in univariate analysis (p < 0.001). However, this relationship did not remain significant when it was examined in the context of P_{plat}, PaO_2/FiO_2, V_T, SAPS, and weight (p = 0.926). In addition, $PEEP$ was not selected for the optimal multivariate model used to predict new ARDS (this model included PIP, sepsis, V_T, and arterial pH). These results suggest that high $PEEP$ alone is not a risk factor for ventilator-associated ALI or ARDS.

5.1.3 Tidal volume (V_T) and normalized tidal volume (V_T/PBW)

There are mixed results regarding tidal volume/normalized tidal volume and their associations with new lung injury. The normalized tidal volume (V_T/PBW) was not associated with development of ALI or ARDS (p = 0.740 and p = 0.559 respectively) and was not considered for multivariate analysis. In addition, the set tidal volume (V_T) was not a risk factor for ALI in patients without lung injury at the outset (p = 0.265). On the other hand, V_T was a significant predictor of ARDS in univariate analysis (p < 0.001) and remained significant in the optimal multivariate model that included PIP, PaO_2/FiO_2, sepsis, pH, and SAPS (p = 0.015). This begs the question, "Why was tidal volume important in some cases and not in others?" The answer partially lies in differences between the patients examined: the ARDS analysis used a larger group of patients who, on average, had sicker lungs at the outset. A larger population increases the statistical power of the analysis. In addition, patients with sicker lungs were more likely to develop further lung injury and may have been ventilated at higher tidal volumes at the outset. The two factors combined may account for the relationship between V_T and development of ARDS. However, when we examine V_T in the context of P_{plat}, patient weight, PaO_2/FiO_2, $PEEP$, and SAPS, this relationship loses significance (p = 0.081). Overall, our results suggest that high tidal volume is associated with ARDS, but the relationship becomes less significant when examined in the context of airway pressures and patient weight. This observation shows the importance of analyzing variables in context of each other using multivariate methods. In this study, high V_T was found to be a significantly associated with ARDS but not ALI. Furthermore, tidal volume was less important compared to P_{plat} and PIP as a risk factor for ARDS.

5.1.4 Results in context of established practice

The finding that ventilator pressures play a greater role in the development of ARDS than tidal volume is supported by existing literature which emphasizes the adverse effects of high airway and transpulmonary pressures [18, 32, 33, 34]. However, this finding also challenges the notion that high tidal volumes in the presence of normal pressures can cause lung injury. Previous studies conclude that tidal volume is the most important risk factor for development of ventilator-induced lung injury [23, 24]. However, these studies did not include patient weight and airway pressure as continuous variables in their multivariate models.

High *PEEP* was also associated with development of ARDS in univariate regression, but this association loses significance in the presence of plateau pressure and PaO_2/FiO_2. Thus it is likely that high *PEEP* is associated with higher plateau pressures and is a marker of sicker patients. Historically, *PEEP* is thought to be lung protective [11, 12] and is often used to recruit collapsed alveoli [35, 36]. However, the results of this study do not point to the protective nature of *PEEP*.

5.1.5 Physiological risk factors

Physiological risk factors for new ARDS include patient obesity, blood acidemia, and high lactate. Obesity decreases respiratory system compliance [37], so obese patients may require higher ventilator pressures to deliver the same tidal volume. The data suggests that patient weight correlates somewhat with tidal volume ($R^2 = 0.19$, $p < 0.001$), and to a lesser extent with set peak inspiratory pressure ($R^2 = 0.05$, $p < 0.001$). Given this information, the exact relationship between weight, tidal volume, and airway pressures remains difficult to discern clearly. In addition, obesity may increase the risk for ARDS through metabolic/systemic effects rather than solely through mechanical effects on the respiratory system. An important randomized clinical trial suggests that obesity may be lung-protective in some cases [10], but our results do not support this hypothesis. Other physiologic risk factors such as low pH and high lactate are characteristic of metabolic acidosis, a condition known to be predictive of acute lung injury in severely traumatized patients [38]. Low pH and high lactate were associated with ARDS, but these associations became less significant in the presence of low PaO_2/FiO_2, high P_{plat}, and high V_T. This suggests that the common practice of using the ventilator to correct for a metabolic acidosis is a potentially harmful intervention if high airway pressures are required. Clinicians are beginning to use permissive hypercapnia as a way to avoid high tidal volumes and high airway pressures in patients with ALI/ARDS [39].

5.1.6 Differences in the patient cohort for analysis of ALI and ARDS

When examining risk factors for the development of ALI and ARDS, we found the associations between day 1 ventilator settings and new ARDS to have more statistical power (lower p-values) compared to those associated with new ALI. For example, *PEEP* was a significant predictor of ARDS ($p < 0.001$) but not of ALI ($p = 0.151$). There are two reasons for these observations. First, the number of patients in the ARDS analysis (789) was larger than that in the ALI analysis (416). A larger group of patients gives more statistical power to the ARDS analysis. Second, the ARDS analysis was based on patients who had sicker lungs at outset of mechanical ventilation compared to that of the ALI analysis. These patients were likely to be given higher tidal volumes/airway pressures and were also predisposed to developing further lung injury. Both reasons contribute to the discrepancies observed between the two patient groups. However, the idea that higher ventilator settings may be an indicator of sicker patients should be addressed by

multivariate models that control for severity of lung and systemic illness.

5.1.7 Recognized limitations

There are several limitations in the current study. First, although data were validated at collection by the bed-side nurses and respiratory therapists, they were not collected by scientific investigators and thus may contain errors. The sheer size of the database should correct for such random errors, assuming inaccuracies are random rather than systematic. Secondly, we assume random variability in initial ventilator settings, but there always exists the possibility that higher pressures and tidal volumes were chosen deliberately to correct underlying hypoxemia, acidemia, and non-cardiogenic pulmonary edema. Knowing this, we have attempted to adjust for severity of illness by including PaO_2/FiO_2, SAPS, and other indicators of underlying illness in the multivariate models. Thirdly, although we used a large patient population in this study, data were not complete in all records. For example, height information was available in 60% of patients, preventing us from calculating tidal volume per predicted body weight for all patients. Missing data reduces the statistical significance of univariate analyses and compels investigators to fill in missing values for multivariate analyses. Finally, PaO_2/FiO_2 values were available only when arterial blood gases were measured, and the accuracy of our patient classifications (no ARDS or ARDS at the outset of ventilator therapy) depended on the presence and validity of these values. In general, any misclassification would bias towards the null hypothesis, making it more difficult to show a relationship between initial ventilator settings and worsening gas exchange in the lungs. Future studies should use pulmonary artery wedge pressure or biomarkers such as BNP in addition to PaO_2/FiO_2 values and chest x-ray reports when diagnosing ARDS. Most importantly, a randomized trial is needed to verify the suggestion that high ventilator pressures play a causal role in the development of ARDS.

5.2 Conclusions

Development of new onset ARDS is a relatively common complication in patients mechanically ventilated $>$ 48 hours in the ICU. High airway pressures, even more than tidal volumes *per se*, are the most important ventilator-associated risk factors for the development of new ARDS. Thus it may be possible to reduce the occurrence of ventilator-induced lung injuries with careful pressure management. However, randomized prospective studies are needed to support this hypothesis. Several physiologic risk factors for ALI/ARDS were identified: these included sepsis, pneumonia, low blood pH, high lactate, and patient obesity. The results of this study contribute to the understanding of ventilator-associated lung injury and the ever changing practice of patient care in the ICU.

5.3 Recommendations for future research

5.3.1 Modifications in the current study

Several modifications of the current study are worth exploring. First, the oxygenation index has been proposed as a better measure of gas exchange in patients who are mechanically ventilated. Thus, it may be useful to identify the patients who have good gas exchange at the outset using this index rather than the PaO_2/FiO_2 ratio. Oxygenation index (OI) is defined as: defined as:

$$OI = \frac{PaO_2 \text{ [Mean airway pressure in cmH}_2\text{O]}}{FiO_2} \tag{5.1}$$

Second, it is possible to perform a sensitivity analysis on criteria used to characterize the patient cohort. For example, how would the classified groups change if patients were required to have 24hrs of healthy gas exchange instead of 12hrs at the outset of mechanical ventilation? How would the results change if we required the deterioration in gas exchange to last 48 hours instead of 24 hours? The data analysis depends heavily on the initial grouping of patients, so it is important to understand how the cohort changes according to inclusion and exclusion criteria.

5.3.2 Related clinical studies

Most groundbreaking studies examine the effects of clinical interventions on patient mortality. Thus, it would be meaningful to study how day one ventilator settings are related to hospital mortality (i.e. use mortality rather than development of ARDS as the primary outcome of interest). Our colleagues at BIDMC also hypothesize that inflammation from injured lungs may spread to the systemic circulation if mechanical ventilation is applied poorly. Thus it is possible to examine the relationship between day one ventilator settings and the development of extra-pulmonary organ failure (such as renal failure).

5.3.3 Other studies using the MIMIC-II database

Many clinicians are interested in how ARDS management has changed since the publication of landmark clinical trials. For example, a large ARDSnet trial showed that low tidal volume ventilation at 6 mL/kg predicted body weight reduced mortality from 39% to 30% compared to traditional 12 mL/kg in patients with ARDS [10]. An important task is to determine whether or not tidal volumes have decreased since the study was published. The MIMIC-II database, which has collected ICU data from 2001 to 2005, is an excellent source of this information. However, the de-identification and date-shifting of all patient records may present a potential obstacle to this particular study.

Appendix A

An Automated Radiology Report Reader

This appendix presents the design and evaluation of the algorithm used to extract information from chest x-ray (text) reports. As described in Chapter 3, the diagnosis of ALI and ARDS depended on finding evidence of bilateral infiltrates in these radiology reports. There were 305 patients with deteriorating gas exchange after the onset of mechanical ventilation, and we needed to examine chest x-ray reports from 12hrs before to 72hrs after the drop in gas exchange for signs of ALI/ARDS. This results in approximately 3.5 reports per patient, which is a total of over 1,000 reports. For this reason, an algorithm was created to help extract information from the text reports and aid in the diagnostic process. The following sections describe the structure and design of the algorithm as well as its performance in a large annotated dataset.

Table A.1: Chest x-ray report for patient b62232 on May 20, 2014 at 12:10 am.

Reason: r/o chf [**Signature 1**]

UNDERLYING MEDICAL CONDITION: 62 year old woman with [**Doctor Last Name 148**]
REASON FOR THIS EXAMINATION: r/o chf [**Signature 1**]

FINAL REPORT

INDICATION: Subarachnoid hemorrhage. Rule out CHF.

COMPARISON: 7 hours earlier.

SINGLE VIEW CHEST: A left subclavian line and right subclavian line are identified. One of the catheters terminates in the mid SVC and the other in the lower SVC. There is no evidence for pneumothorax. There is increased opacity of both lungs, greater on the right than on the left. This may be consistent with pulmonary edema. There is no evidence for congestive heart failure. No pleural effusions are present.

IMPRESSION: Findings consistent with pulmonary edema.

A.1 An example chest x-ray report

Table A.1 shows an x-ray report for a patient who developed ARDS after the onset of mechanical ventilation. Each radiology report contains information such as the time of exam, the purpose of the exam, and clinical interpretations from the radiologist. On average, x-rays were taken approximately once per day and reports were recorded at the same rate. Note that this report was divided into multiple sections: underlying conditions, reason for examination, findings, impression, etc. The clinical interpretation of the x-ray film was included in sections entitled SINGLE VIEW CHEST, and IMPRESSION.

A.2 Algorithm design

Figure A.1 shows the main components of the text analysis algorithm. These components included (1) a **report parser**, which reads the report and identifies sections that contain clinical interpretations/diagnoses; (2) a **search engine**, which examines the relevant parts of the report for specific phrases (ex. *opacities, infiltrates*); and (3) a **logical interpreter** that uses the results of the search engine to produce desired outputs (ex. presence or absence of bilateral infiltrates). A key advantage of this algorithm is the ability to change search phrases and output rules without altering the algorithm itself. The three main components are further described in the following sections.

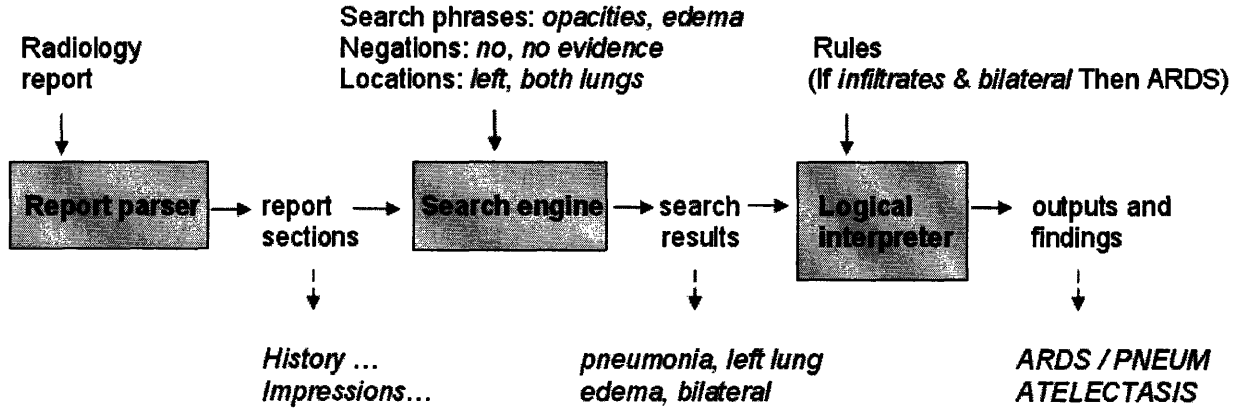


Figure A.1: A schematic of components in the text-analysis algorithm.

A.2.1 The report parser

The report parser prepares a report for the search engine in the following manner. First, it removes extra spaces from the beginning and end of each line. In the original report, these spaces were used as tabs to align the report, possibly for printing purposes. Second, the parser identifies the beginning of each section by looking for a colon (:) at the start of a paragraph. The text before the colon is used as the title of the section, and finding a new colon terminates the previous section. For example, "COMPARISON: 7 hours earlier" would be considered one section with "COMPARISON" as the title. Third, the parser labels sections that appear after "FINAL REPORT." This prevents the algorithm from searching in undesired parts of the report, for example in the sections corresponding to past medical history (i.e. "UNDERLYING MEDICAL CONDITION"). Table A.2 lists the sections and properties extracted from the example x-ray report.

Table A.2: Report sections and properties for the example chest x-ray report.

Section title	**Part of "FINAL REPORT"**
Reason	no
UNDERLYING MEDICAL CONDITION	no
REASON FOR THIS EXAMINATION	no
INDICATION	yes
COMPARISON	yes
SINGLE VIEW CHEST	yes
IMPRESSION	yes

A.2.2 The search engine

The search engine is the core of the text analysis algorithm. It is responsible for finding specific phrases in a report, identifying negations and locations associated with the phrase, and returning results to the logic interpreter. The engine examines all parts of the radiology report that appear after "FINAL REPORT," excluding sections that describe previous illnesses ("UNDERLYING ILLNESS", "HISTORY", "INDICATION", "ADMITTING", etc). For a given search phrase, the engine examines the report in the following manner:

1. Determine if a section contains a the particular phrase via a normal linear search.
2. If the phrase is present, find the specific sentence that contains the phrase. Examine this sentence for:
   - **Negations**, or words that indicate the item is not present (ex. "there is *no evidence* of edema"). A list of negations was obtained from the NegEx algorithm, which was originally designed to extract diseases from medical discharge summaries [40]. Two different word lists were used to identify

negations before and after the search phrase; these lists are shown in Tables A.3 (post-phrase negations) and A.4 (pre-phrase negations).

- **Physical location**, or location in the human body (ex. infiltrates are present in *both lungs* and atelectasis is seen at the *left base*). A list of terms used to identify physical locations is shown in Table A.5. The algorithm looked for words that describe physical location in the same sentence as the search phrase.
- **Phrase descriptions**, or words that further characterize the disease (ex. there is a *diffuse* alveolar pattern in the lungs). A list of search phrases and associated descriptions searched by the algorithm is shown in Table A.6.

3. Return the search results to the logic interpreter for further processing. The process is then repeated for other search phrases. In the example report, the algorithm found the following diseases: *OPACITY: both lungs, EFFUSIONS: no, and EDEMA: pulmonary.*

Table A.3: Post-phrase negations, modified from NegEx [40]

Negation terms
unlikely
free
was ruled out
is ruled out
are ruled out
have been ruled out
has been ruled out
is not present

Table A.4: Pre-phrase negations, modified from NegEx [40]

Negation terms		
absence of	not have	did rule the patient out
cannot	not know of	did rule him out for
cannot see	not known to have	did rule her out for
checked for	not reveal	did rule him out against
declined	not see	did rule her out against
declines	not to be	did rule the patient out for
denied	patient was not	did rule the patient out against
denies	rather than	can rule out
denying	resolved	can rule out for
evaluate for	test for	can rule out against
fails to reveal	to exclude	can rule him out
free of	unremarkable for	can rule her out
negative for	with no	can rule the patient out
neither	without	can rule him out for
never developed	without any evidence of	can rule her out for
never had	without evidence	can rule the patient out for
no	without indication of	can rule him out against
no abnormal	without sign of	can rule her out against
no cause of	rules out	can rule the patient out against
no complaints of	rules him out	adequate to rule out
no evidence	rules her out	adequate to rule him out
no new evidence	rules the patient out	adequate to rule her out
no other evidence	rules out for	adequate to rule the patient out
no evidence to suggest	rules him out for	adequate to rule out for
no findings of	rules her out for	adequate to rule him out for
no findings to indicate	rules the patient out for	adequate to rule her out for
no mammographic evidence of	ruled out	adequate to rule the patient out for
no new	ruled him out	adequate to rule the patient out against
no radiographic evidence of	ruled her out	sufficient to rule out
no sign of	ruled the patient out	sufficient to rule him out
no significant	ruled out for	sufficient to rule her out
no signs of	ruled him out for	sufficient to rule the patient out
no suggestion of	ruled her out for	sufficient to rule out for
no suspicious	ruled the patient out for	sufficient to rule him out for
not	ruled out against	sufficient to rule her out for
not appear	ruled him out against	sufficient to rule the patient out for
not appreciate	ruled her out against	sufficient to rule out against
not associated with	ruled the patient out against	sufficient to rule him out against
not complain of	did rule out	sufficient to rule her out against
not demonstrate	did rule out for	sufficient to rule the patient out against
not exhibit	did rule out against	resolution of
not feel	did rule him out	
not had	did rule her out	

Table A.5: Terms that describe physical location in the lungs

Location	Terms
Both lungs	bilateral, bilaterally, bibasilar, bilat, both lungs, perihilar, multifocal, both lobes, both lung zones, both apices, both bases, lower lung zones, upper lung zones
Left lung	left, retrocardiac, behind the heart, lingula
Right lung	right
Left or right lung	asymmetric, hemithorax
Not in lungs	abdomen, artery, breast

Table A.6: Phrases and associated descriptions searched by the text analysis algorithm.

Phrases	Descriptions
respiratory distress, rds, ards	
pneumonia	
atelectasis, atelectases, atelectatic	
collapse	
effusion(s)	pleural
fluid	pleural
consolidation(s)	diffuse, hazy, patchy, opaque
disease	diffuse, hazy, patchy, opaque
infiltrate(s)	diffuse, hazy, patchy, opaque
density(ies)	diffuse, hazy, patchy, opaque
alveolar pattern	diffuse, hazy, patchy, opaque
fullness	diffuse, hazy, patchy, opaque
haziness, haze	diffuse, patchy, opaque
opacity(ies), opacified, opacification(s)	diffuse, hazy, patchy
edema	pulmonary, pulm, interstitial
lung(s)	inflammatory, inflammation, inflammed, clear

A.2.3 The logical interpreter

The logical interpreter examines the output of the search engine and makes diagnoses based on a list of rules. An example rule is: if both *bilateral* and *infiltrates* are present, mark the patient as having *bilateral infiltrates*. The logical interpreter loads the set of rules from an easily manipulated medium, such as a text file. It then examines the results returned by the search engine to see if any rules are satisfied. If so, the appropriate diagnoses are recorded. The list of rules used by the interpreter is shown in Tables A.7 and A.8. In the example report, the patient was identified as having BILAT-INFILTRATES.

Table A.7: Rules used to extract *findings* from search results

Finding	Criteria
infiltrates	ARDS, inflammation, inflammatory, inflammed, pneumonia, patchy, infiltrate(s), density(ies), edema, hazy(iness), opacity(ies), opacification, opacified, respiratory distress, diffuse, or fullness
atelectasis	atelectasis, atelectases, atelectatic, collapse, consolidation, or consolidations
effusion	effusion(s)
clear	clear
bilateral	bilateral(ly), bibasilar, apices, bases, lower lung zones, or upper lung zones
bilateral	both + (lung zones, lobes, lungs, apices, or bases)
bilateral	left + right
bilateral	NOT (left or right) + (perihilar, multifocal, interstitial, pulmonary, pulm)
right	right
left	left, retrocardiac, behind the heart
lungs	NOT (abdomen, artery, or breast)

A.3 Algorithm performance

The text analysis algorithm was evaluated by a manual review of chest x-ray reports in 305 patients with deteriorating gas exchange. A preliminary version of the algorithm was used to detect bilateral infiltrates and atelectasis/consolidations in these patients. In patients where bilateral disease was found, one report that contained the diagnosis was selected and examined manually. In all other patients, reports from 12hrs before the drop in gas exchange to 72hrs after were reviewed. In total, 641 reports were examined for the presence of infiltrates, atelectasis, pleural effusions, and clear lungs. Distinctions were made between diseases present in the left, right, or bilateral lungs. If a particular disease was present in both the left and right lungs, it was considered a bilateral disease. Table A.9 shows the incidence of each disease in this gold standard of 641 annotated reports.

The output of the text-analysis algorithm was then compared to the gold standard. The sensitivity, positive

Table A.8: Rules used to extract DIAGNOSES from *findings*

Diagnosis	**Finding 1**	**Finding 2**	**Finding 3**
BILAT-INFILTRATES	*infiltrates*	*bilateral*	*lungs*
L-INFILTRATES	*infiltrates*	*left*	*lungs*
R-INFILTRATES	*infiltrates*	*right*	*lungs*
INFILTRATES	*infiltrates*	*NOT (left, right, or bilateral)*	*lungs*
BILAT-ATELECTASIS	*atelectasis*	*bilateral*	*lungs*
L-ATELECTASIS	*atelectasis*	*left*	*lungs*
R-ATELECTASIS	*atelectasis*	*right*	*lungs*
ATELECTASIS	*atelectasis*	*NOT (left, right, or bilateral)*	*lungs*
BILAT-EFFUSION	*effusion*	*bilateral*	*lungs*
L-EFFUSION	*effusion*	*left*	*lungs*
R-EFFUSION	*effusion*	*right*	*lungs*
EFFUSION	*effusion*	*NOT (left, right, or bilateral)*	*lungs*
BILAT-CLEAR	*clear*	*bilateral*	*lungs*
L-CLEAR	*clear*	*left*	*lungs*
R-CLEAR	*clear*	*right*	*lungs*
CLEAR	*clear*	*NOT (left, right, or bilateral)*	*lungs*

Table A.9: Findings from a manual review of 641 reports.

	Bilateral lungs	Left lung	Right lung	Not present
Infiltrates	256	89	41	255
Atelectasis	162	168	42	269
Effusions	154	89	58	340
Clear	31	8	17	585

predictive value (PPV), and accuracy was calculated for each disease. The definitions for sensitivity, PPV, and accuracy are the follows:

$$Sensitivity = \frac{\#true\ positives}{\#true\ positives\ +\ \#false\ negatives}$$

$$Positive\ predictive\ value = \frac{\#true\ positives}{\#true\ positives\ +\ \#false\ positives}$$

$$Accuracy = \frac{\#true\ positives\ +\ \#true\ negatives}{\#total\ records}$$

In this thesis, a *true positive* was defined as a correct prediction of a disease or characteristic that was also associated with the correct part of the lungs. A *false positive* was a positive prediction that did not result in a *true positive*. *False negatives* and *true negatives* assume their normal definitions. The performance of the algorithm in terms of sensitivity, PPV, and accuracy are listed in Table A.10.

Table A.10: Performance of the radiology report analysis algorithm.

	Infiltrates	**Atelectasis**	**Effusions**	**Clear**
n (out of 641)	386	372	301	56
Sensitivity	0.98	0.98	0.98	0.96
PPV	0.95	0.92	0.95	0.98
Accuracy	0.96	0.94	0.97	0.996

In general, the algorithm had high sensitivity ($\geq$ 0.97), high positive predictive value ($\geq$ 0.92), and high accuracy ($\geq$ 0.94) among the different types of diseases/characteristics extracted. The differences between the gold standard and the algorithm predictions were reviewed, and the following observations were made. First, the algorithm was in general more accurate than the human reader because it could systematically identify every instance of the disease. Second, the most common error made by the algorithm was the inability to differentiate between multiple physical locations in the same sentence. For example, "there are infiltrates in the left and effusions in the right lung," was interpreted as bilateral infiltrates and bilateral effusions. Addressing this issue is non-trivial and may be part of future work that expand on the current algorithm.

The automated text analysis algorithm runs fairly quickly: it can evaluate reports at 1,000 patients per minute (where each patient has on the order of 10 reports). At its current speed and accuracy, it may prove a useful tool for identifying patients from the MIMIC-II database who have specific respiratory diseases. However, a manual review of the identified patients is still recommended to ensure correct diagnoses.

A.4 Algorithm code

The "RadiologyReader" package was build using Netbeans IDE 5.5 and Java 1.5. The list of java classes that make up this package are shown and described in Table A.11.

Table A.11: Source files from the RadiologyReader Package

File	Description
Main.java	main class to run RadiologyReader
RadiologyParser.java	parses a radiology report into sections
RadiologySearchEngine.java	main search engine component of the algorithm
RadiologyWriter.java	writes the findings/diagnoses to a text file
RadiologyFinding.java	class for one particular finding/diagnoses
RadiologyFindings.java	class for a set of findings/diagnoses
ReportList.java	class for a list of radiology reports
ReportSection.java	class for a particular section of the radiology report
Rule.java	class for a particular rule used in by logical interpreter
RuleReader.java	reads user-written rules from a text file
SearchResult.java	class for the result of an algorithm search
SearchResults.java	class for a list of results returned by the search
SearchTerm.java	class to contain a searched phrase and modifier terms
SingleReport.java	class for one radiology report
WordList.java	container of a list of words used by the algorithm

Bibliography

[1] Massachusetts General Hospital Cancer Resource Room. http://www.massgeneral.org/cancer/crr/types/thoracic/illustrations/lungs_front.asp.

[2] GR Bernard, A Artigas, and KL Brigham et al. Report of the American-European consensus conference on ARDS: definitions, mechanisms, relevant outcomes and clinical trial coordination. *Am Rev Respir Crit Care Med*, 149:818–824, 1994.

[3] PL Maurino. *The ICU Book.* Williams & Wilkins, Baltimore, MD, 1998.

[4] JJ Marini. Evolving concepts in ventilatory management of acute respiratory distress syndrome. *Clin Chest Med*, 17:555–575, 1996.

[5] L Gattinoni and A Pesenti. The concept of "baby lung". *Intensive Care Med*, 31:776–784, 2005.

[6] MB Amato, CS Barbas, and DM Medeiros et al. Effect of a protective-ventilation strategy on mortality in the acute respiratory distress syndrome. *N Engl J Med*, 338:347–354, 1998.

[7] TW Stewart, MO Meade, and DJ Cook et al. Evaluation of a ventilation strategy to prevent barotrauma in patients at high risk for acute respiratory distress syndrome: Pressure- and volume-limited ventilation strategy group. *N Engl J Med*, 338:355–361, 1998.

[8] L Brochard, F Roudot-Thoraval, and E Roupie et al. Tidal volume reduction for prevention of ventilator-induced lung injury in acute respiratory distress syndrome. *Am J Respir Crit Care Med*, 158:1831–1838, 1998.

[9] RG Brower, CB Shanholtz, and HE Fessler et al. Prospective, randomized, controlled clinical trial comparing traditional versus reduced tidal volume ventilation in acute respiratory distress syndrome patients. *Crit Care Med*, 27:1492–1498, 1999.

[10] The Acute Respiratory Distress Syndrome Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med*, 342:1301–1308, 2000.

[11] JG Muscedere, JB Mulle, and K Gan et al. Tidal ventilation at low airway pressures can augment lung injury. *Am J Respir Crit Care Med*, 149:1327–1334, 1994.

[12] D Dreyfuss, P Soler, and G Basset et al. High inflation pressure pulmonary edema. respective effects of high airway pressure, high tidal volume, and positive end-expiratory pressure. *Am J Respir Crit Care Med*, 137, 1988.

[13] The National Heart Lung and Blood Institute ARDS Clinical Trials Network. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *N Eng J Med*, 351:327–336, 2004.

[14] GM Saul, TW Freeley, and FG Mihm. Effect of graded PEEP on lung water in noncardiogenic pulmonary edema. *Crit Care Med*, 10:667–669, 1982.

[15] C Helbert, A Paskanik, and CE Bredenberg. Effect of positive end-expiratory pressure on lung water in pulmonary edema caused by increased membrane permeability. *Ann Thorac Surg*, 311:281–286, 1984.

[16] T Kolobow, MP Moretti, and R Fumagalli et al. Severe impairment in lung function induced by high peak airway pressure during mechanical ventilation: an experimental study. *Am Rev Respir Dis*, 135:312–315, 1987.

[17] D Dreyfuss, G Basset, and P Soler et al. Intermittent positive-pressure hyperventilation with high inflation pressures produces pulmonary microvascular injury in rats. *Am Rev Respir Dis*, 132:880–884, 1985.

[18] OU Gurkan, C O'Donnell, and R Brower at al. Differential effects of mechanical ventilatory strategy on lung injury and systemic organ inflammation in mice. *Am J Physiol Lung Cell Mol Physiol*, 2003.

[19] D Dreyfuss, P Soler, and G Saumon. Mechanical ventilation-induced pulmonary edema. interaction with previous lung alterations. *Am J Respir Crit Care Med*, 151, 1995.

[20] A Esteban, A Anzueto, and F Frutos at al. Characteristics and outcomes in adult patients receiving mechanical ventilation. *J Am Med Assoc*, 287:345–355, 2002.

[21] A Carlucci, JC Richard, and M Wysocki et al. Noninvasive versus conventional mechanical ventilation. *Am J Respir Crit Care Med*, 163, 2001.

[22] RA Garibaldi, MR Britt, and ML Coleman et al. Risk factors for postoperative pneumonia. *Am J Med*, 70, 1981.

[23] O Gajic, SI Dara, and JL Mendez et al. Ventilator-associated lung injury in patients without acute lung injury at the onset of mechanical ventilation. *Crit Care Med*, 32, 2004.

[24] O Gajic, F Frutos-Vivar, and A Esteban et al. Ventilator settings as a risk factor for acute respiratory distress syndrome in mechanically ventilated patients. *Intensive Care Med*, 31, 2005.

[25] M Saeed, C Lieu, and G Raber et al. Mimic ii: a massive temporal icu patient database to support research in intelligent patient monitoring. *Computers in Cardiology*, 29:641–644, 2002.

[26] DS Lee, C Tran, and V Flintoft et al. Ccort/ccs quality indicators for congestive heart failure care. *Can J Cardiol*, 19, 2003.

[27] HM Krumholz, DW Baker, and CM Ashton et al. Evaluating quality of care for patients with heart failure. *Circulation*, 101:e122, 2000.

[28] FH Ramadan, N Masoodi, and AA El-Solh. Clinical factors associated with hyperkalemia in patients with congestive heart failure. *J Clin Pharm Therap*, 30:223–239, 2005.

[29] JM Bland and DG Altman. The odds ratio. *BMJ*, 320:1468, 2000.

[30] JAC Sterne and GD Smith. Sifting the evidence–what's wrong with significance tests? *BMJ*, 322, 2001.

[31] RA Fisher. *Statistical methods for research workers.* London: Oliver and Boyd, 1950.

[32] D Talmor, T Sarge, and CR O'Donnell et al. Esophageal and transpulmonary pressures in acute respiratory failure. *Crit Care Med*, 34, 2006.

[33] K Tsuno, P Prato, and T Kolobow. Acute lung injury from mechanical ventilation at moderately high airway pressures. *J Applied Physiology*, 69, 1990.

[34] MH Cho, A Malhotra, and DM Donahue et al. Mechanical ventilation and air leaks after lung biopsy for acute respiratory distress syndrome. *Ann Thorac Surg*, 82:261–266, 2006.

[35] CP Wu, KL Huang, and WC Perng et al. Optimization of ventilation-perfusion matching during protective lung ventilation. *Chest*, 128, 2005.

[36] R Baba, K Paramesh, and JW Zwall et al. Protective ventilation of patients with acute respiratory distress syndrome. *British Journal of Anaesthesia*, 93, 2004.

[37] J Sprung, DG Whalley, and T Falcone et al. The impact of morbid obesity, pneumoperitoneum, and posture on respiratory system mechanics and oxygenation during laparoscopy. *Anesth Analg*, 94:1345–1350, 2002.

[38] LW Eberhard, DJ Morabito, and MA Matthay et al. Initial severity of metabolic acidosis predicts the development of acute lung injury in severely traumatized patients. *Crit Care Med*, 28, 2000.

[39] DA Kregenow and ER Swenson. The lung and carbon dioxide: implications for permissive and therapeutic hypercapnia. *Eur Respir J*, 20:6–11, 2002.

[40] WW Chapman, W Bridewell, and P Hanbury et al. A simple algorithm for identifying negated findings and diseases in discharge summaries. *J Biomed Inform*, 34:301–10, 2001.