The Index of Respiratory Distress:

A Clinical Decision Rule to Assist Out-of-Hospital Providers in Caring for Older Patients in Respiratory Distress

by

Susan E. Shapiro

A Dissertation

Presented to
Oregon Health & Science University
School of Nursing
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

August 21, 2003

APPROVED:

[redacted]

Linda McCauley, PhD, RN, FAAN, Scientist and Professor, Center for Research on Occupations and Environmental Toxicology and School of Nursing
Research Advisor

[redacted]

Mohamud Daya, MD, FACEP, Associate Professor, School of Medicine,
Committee Member

[redacted]

Janet Larsen, PhD, RN, FAAN, Professor, College of Nursing, University of Illinois at Chicago, Committee Member

[redacted]

Anne Rosenfeld, PhD, RN, Associate Professor, School of Nursing, Committee Member

[redacted]

Kathleen Potempa, DNSc, RN, FAAN, Dean, School of Nursing

Acknowledgment

This project would not have been possible without the loving support of my family and friends. To my Portland family – both the "real" family and those who became family during this process, I am grateful for your encouragement and faith in my work. To my two brothers, both of whom are distant from here, and the distant friends as well, I need to say a special thanks; you are never far from my thoughts or my heart. Your collective faith in me has been inspiring.

A very special "thank you" goes to Michael Lasarev, statistician *extraordinaire* - and special friend - for his patience in helping me appreciate the subtleties of data analysis and interpretation. There is truly a beauty in statistics well done, and he is a pro; I owe him big-time.

I'm sure others think they've had the best dissertation advisor there is, but they're wrong; I've had her. I want to thank Dr. McCauley for stepping up and offering her assistance, and then for providing the kind of opportunities and guidance that every graduate student wishes for. She has been the perfect advisor, there when I needed her and providing firm, but gentle guidance as I learned the craft of research. I can't thank her enough. The remaining members of my committee have also been great; thank you all for your willingness to support me in this project; your insights have been extremely valuable.

I also want to take this opportunity to thank the out-of-hospital providers and support personnel who gave so ~~generous~~ generously of their time and expertise to assist me in my

efforts. Similarly, a heartfelt thank you to those who helped identify cases for the data abstraction phase, and the medical records personnel at the receiving hospitals who helped pull charts for review.

On a much more personal note, there are some who made this possible who are no longer here. Completing this doctorate would have given both my parents, Joseph and Judith Shapiro, great pleasure. They provided me with the necessary raw material to face all that life has sent my way and I'm only sorry I couldn't manage to do this while they were still alive to enjoy it. And finally, but most importantly, I want to acknowledge with great love and tenderness, my late husband Gary Michael Allums. He made this entire journey possible in more ways than I can express and I know he would have delighted in its completion. It is with a mixture of profound joy and deep sadness, then, that I lovingly dedicate this dissertation to his memory.

ABSTRACT

TITLE: The Index of Respiratory Distress: A Clinical Decision Rule to Assist Out-of-Hospital Providers in Caring for Older Patients in Respiratory Distress

AUTHOR: Susan E. Shapiro

APPROVED: [redacted]
Linda McCauley, PhD, RN, FAAN

The purpose of this project was to begin developing a clinical decision rule (CDR) for evaluating respiratory distress (RD) in older out-of-hospital (OOH) patients. Specifically, the study aimed to: 1) determine the appropriate items, scaling, and scoring of the instrument and 2) assess the instrument's reliability and validity. The investigator developed a definition of RD, then reviewed the medical, nursing, and out-of-hospital literature to determine the clinical indicators that have been described as indicating various levels of severity of respiratory distress. This list was reduced to seven assessment parameter using a Delphi technique to obtain input from six experts in out-of-hospital care.

Data on the seven predictors were abstracted from the OOH medical records of 124 cases of RD patients 50 years of age and older. Severity was determined from review of the emergency department (ED) record matched to the OOH event. A binary logistic regression was done, with "severe distress" as the outcome of interest. Accessory muscle use, abnormal respiratory effort, and the inability to speak full sentences were identified as the three best predictors of severe respiratory distress in older OOH patients. However,

these results were difficult to interpret primarily due to a relatively small sample size and problems encountered in coding the predictor variables.

Despite the numerous limitations cited, the results of this study demonstrate the possibility of determining severe RD from clinical assessment parameters currently being used by OOH providers. This suggests that with further work, it will be possible to develop an Index of Respiratory Distress that can assist OOH providers in caring for older patients in respiratory distress.

TABLE OF CONTENTS

CONTENTS PAGE

CONTENTS PAGE

CONTENTS PAGE

CONTENTS PAGE

CONTENTS PAGE

Shapiro, Susan.
The Index of Respiratory Distress: A Clinical Decision Rule to Assist Out-of-Hospital Providers in Caring for Older Patients in Respiratory Distress

ERRATA

Page	Current Content	Correct Content
iii	bottom of page, *generous*	*generously*
xiii-xiv	List of Tables: pages in error	See below
100	Table 15, Reason for 9-1-1 call, not respiratory reads *66(21.2)*	Should read *10(3.2)*

Corrected List of Tables:

LIST OF TABLES

PAGE

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	TERM
ACLS	Advance Cardiac Life Support
ALOC	Altered Level of Consciousness
ALS	Advanced Life Support
AVPU	Awake; responds to Verbal stimulus only; responds to Painful stimulus only; Unconscious
BiPAP	Bi-level Positive Airway Pressure
BLS	Basic Life Support
BTF	Brain Trauma Foundation
BVM	Bag-valve-mask
CA	Cardiac Arrest
CDR	Clinical Decision Rule
CHF	Congestive Heart Failure
COPD	Chronic Obstructive Pulmonary Disease
CPAP	Continuous Positive Airway Pressure
DDI	Dyspnea Differentiation Index
DLOF	Deviance Lack of Fit
ED	Emergency Department
EMS	Emergency Medical Services
EMSOP	Emergency Medical Services Outcome Project
EMT	Emergency Medical Technician
ETI	Endotracheal Intubation
FWA	Federalwide Assurance
GCS	Glasgow Coma Scale

ABBREVIATION	TERM
HL	Hosmer-Lemeshow lack of fit
IRB	Institutional Review Board
IRD	Index of Respiratory Distress
ISS	Injury Severity Score
LOC	Level of Consciousness
MI	Myocardial Infarction
MVC	Motor Vehicle Crash
NHLBI	National Heart, Lung, and Blood Institute
NHTSA	National Highway Traffic Safety Administration
OHSU	Oregon Health & Science University
OOH	Out-of-Hospital
OOH CA	Out-of-Hospital Cardiac Arrest
OPALS	Ontario Prehospital Advanced Life Support
PaO_2	Arterial Partial Pressure of Oxygen
PEFR	Peak Expiratory Flow Rate
PHI	Protected Health Information
PSAP	Public Safety Answering Point
RAM	Risk Adjustment Measure
RCT	Randomized Clinical Trial
RD	Respiratory Distress
SaO_2	Oxygen Saturation

CHAPTER 1

INTRODUCTION

Specific Aims

The purpose of this project is to begin developing a clinical decision rule (CDR) for evaluating respiratory distress (RD) in out-of-hospital patients.[1] Clinical decision rules are tools designed to assist providers in clinical decision-making. They consist of three or more predictor variables derived from the patient's history, physical examination, or laboratory findings, and an outcome variable that is either the probability of a diagnostic outcome or a prescribed treatment. The rule is created using multivariate analytic techniques as opposed to a consensus-building process (Laupacis, Sekar, & Stiell, 1997; McGinn et al., 2000; Stiell & Wells, 1999; Wasson, Sox, Neff, & Goldman, 1985).

The proposed Index of Respiratory Distress (IRD) would prescribe a standard assessment for OOH patients in respiratory distress, and identify those in severe distress. For patients identified in severe distress, OOH providers would institute appropriate advanced life support (ALS) measures; for those not in severe distress, providers would

[1] The language in this area, like much of the language of health care, has changed in recent years. What was once known as "prehospital care," that is, EMS responses to community based emergency calls (usually designated as 9-1-1 responses) has been re-named "out-of-hospital care." This was done because EMS-type vehicles and personnel are being used more and more frequently for non-9-1-1 responses such as interfacility transfers. For the purposes of this project, however, out-of-hospital care refers only to 9-1-1 responses.

provide basic life support (BLS), reassurance, and safe transport to the emergency department for definitive diagnosis and care.[2]

In addition to predicting severe respiratory distress, the IRD will make four significant contributions to the clinical care of older OOH patients in respiratory distress. First, it will help standardize the language used to describe respiratory distress. Second, it will help standardize the assessment of RD patients in the OOH setting, thus allowing for more consistent assessments of patients' responses to interventions. These first two contributions will result in the third: improved communications among all levels of OOH providers, and between the OOH providers and emergency department nurses and physicians. Finally, the IRD may help reduce the incidence of error in treating older adult patients in respiratory distress by limiting ALS interventions to patients in severe distress.

The IRD will also contribute to research in the OOH setting. By standardizing the assessment of OOH patients in respiratory distress, it will support evaluations of various ALS interventions within and across agencies. It will also standardize the documentation across agencies, allowing for easier comparisons of treatment protocols and their attendant outcomes.

Developing, validating, and disseminating CDRs is a complex, long-term undertaking. This proposal addresses only the initial steps in developing the Index of

---

[2] Basic Life Support and Advanced Life Support refer to the different skill levels of different levels of OOH providers. In general, basic Emergency Medical Technicians (EMTs) are limited to BLS skills such as splinting, bandaging, spinal immobilzation, and administration of oxygen. Advanced life support skills are performed by OOH providers with more training (including nurses), and include, for example, endotracheal intubation, venous cannulation, and administration of various medications. The actual scope

Respiratory Distress. Specifically, the study aims to: 1) determine the appropriate items, scaling, and scoring of the instrument and 2) assess the instrument's reliability and validity.

Background and Significance

Outcomes of out-of-hospital care

The Emergency Medical Services (EMS) systems currently in operation in most of the United States can trace their existence to the Emergency Medical Services Act of 1973, although there was important scientific work that preceded the act (National Highway Traffic Safety Administration (NHTSA) 1996; Howard, 2000; Rose, 1980). These systems arose in response to two parallel medical developments. One was improvement in the care of soldiers injured in battle; these improvements were first noted in Korea in the early 1950's, and continued in Vietnam in the late 1960's and early 1970's. The second was the emergence of intensive care and resuscitation for victims of myocardial infarctions (MIs) and sudden cardiac death (NHTSA, 1996; Howard; Spaite, Criss, Valenzuela, & Meislin, 1997). While much of the focus of EMS continues to be on care of trauma patients and those with acute myocardial infarction/sudden cardiac death, the scope of care in the OOH setting has expanded significantly, and providers treat patients with a wide variety of problems and varying levels of acuity.

Early authors identified several challenges to EMS evaluation and research. These included the lack of a system focus for the research (Gibson, 1973), the complexity of the

---

of practice of each level of provider and the skills they may perform is determined by states and local jurisdictions and vary considerably from place to place.

EMS system itself (Rose), and the lack of reliable performance measures (Rose). Initial evaluations of EMS systems focused on structural components, particularly response times of ambulances and times to administration of selected interventions (Abramson & Safar, 1990; Gibson, 1973). Although even early authors identified problems with this approach, published evaluation and research reports reflect its continued use (Abramson & Safar; Narad & Driesbock, 1999). The lack of research focusing on the outcomes of OOH interventions has been well-documented (Brice, Garrison, & Evans, 2000; Callaham, 1997; Maio et al., 1999; Spaite, Criss, Valenzuela, & Guisto, 1995); in fact, there is actually very little known about the effectiveness of most standard OOH ALS interventions.

Setting Priorities for Out-of-Hospital Research

Recognizing the need for more comprehensive and better designed studies of outcomes of OOH care, the National Highway Traffic Safety Administration (NHTSA) funded, a multi-million dollar, multi-site study called the Emergency Medical Services Outcomes Project (EMSOP) (Maio et al.). The project had three objectives: the first, to identify priority conditions for EMS research; the second, to identify risk adjustment measures for these conditions; and the third, to identify appropriate outcome measures.

Results of the first stage of the project were reported by Maio, et al. in 1999. In order to identify the priority conditions, the research group used a large EMS data set to determine the most frequently encountered conditions. A total of over 391,000 cases were analyzed, representing over 13 million people of all ages from various regions throughout

the country. Using a panel of EMS experts, including physicians, nurses, and paramedics, the group identified the top medical and surgical conditions for which OOH interventions would have the greatest impact on any of 6 identified outcomes. The outcomes of interest were known as "the six Ds," death (survival), disease (impaired physiology), disability (limiting potential disability), discomfort (alleviating discomfort), dissatisfaction (satisfaction), and destitution (cost-effectiveness) (Maio, et al.,p. 425).

This process presented special challenges, since as was indicated above, very little is known about how any OOH interventions affect any of the identified outcomes, other than survival from OOH CA. Therefore, the investigators in the EMSOP study were forced to rely on the consensus of experts (Maio et al.). Acknowledging this constraint, they derived a weighted score for each condition, accounting for its frequency, the extent to which OOH interventions may be relevant to each of the six outcomes, and the extent to which OOH interventions were likely to positively affect each outcome. Seven conditions were identified in the top quartile for both adults and children; respiratory distress was identified as the third priority condition for OOH outcomes research for children (after major and minor trauma) and the second priority condition for adults (after minor trauma). For both adults and children, respiratory distress was identified as the top priority **medical** condition for OOH outcomes research (Maio et al.).

Out-of-Hospital Care of Patients in Respiratory Distress

Patients present to OOH providers in varying degrees of RD and with varying levels of physiologic compromise. In addition, respiratory distress may be indicative of a wide variety of illnesses or injuries that may require a range of medical or surgical interventions. Underlying causes may range from those that are relatively benign to those that are immediately life- threatening. Persons of any age, from neonates to the very elderly, may present with complaints or indications of respiratory distress.

Elderly patients in RD provide special challenges for OOH providers. Among elderly patients, it is often difficult to differentiate pulmonary causes of RD such as chronic obstructive pulmonary disease (COPD) or pneumonia, from RD caused by congestive heart failure (CHF) (Ailani et al., 1999; Mosesso, Dunford, Blackwell, & Griswell, 2003). Accurately determining the cause of acute RD in older adults is a challenge even in the controlled environment of an emergency department, where auscultation of lung sounds can be carried out in a relatively stable setting and chest radiography is readily available. The OOH setting lacks both opportunities.

As Ailani and colleagues (1999) pointed out, accurate diagnosis of these patients is crucial because the treatments for RD of cardiac origin are quite different than those for RD of pulmonary origin, and erroneous treatment may prove harmful. In an oft-cited study of OOH patients in RD, Weurz & Meador (1992) looked at mortality and hospital length of stay for RD patients treated in the OOH setting for CHF. Their study showed that 15% of those patients were, in fact, misdiagnosed, and as a result, treated

erroneously. They grouped the patients in their study as critical or non-critical based on subjective paramedic assessment, acknowledging the fact that no generally accepted severity scales existed. What they observed was that for patients judged to be critical, treatment for CHF with standard ALS medications improved mortality, *even among those who were treated erroneously.* For those who were judged not critically ill, patients treated erroneously for CHF showed significantly increased mortality. In neither case did ALS treatment have any effect on hospital length of stay.

The study by Weurz & Meador (1992) was a retrospective case series analysis, and suffers from the biases inherent in such studies, including selection bias. The authors acknowledged this problem, along with other limitations. Nonetheless, their results suggest that ALS treatment for patients in CHF be limited to those in severe distress, where errors in diagnosis and treatment seem to have less of an adverse impact on at least one important outcome measure, survival. It would be virtually impossible, however, to replicate their study since they did not use any standard to define levels of respiratory distress other than paramedic judgment.

Defining and Measuring Respiratory Distress

Defining respiratory distress

Although the concept of RD is used widely in medical, nursing, and prehospital literature, it is rarely defined, and not consistently described. A review of selected general medical and nursing texts (Craven & Hirnle, 2000; Davies & Hoffman, 2000; Reinke & Hoffman, 2000; Thelan, Lough, Urden, & Stacy, 1998; Thompson, McFarland,

Hirsch, & Tucker, 1997), texts focusing on pulmonary or respiratory care (Cronin, 1997; Dantzker, MacIntyre, & Bakow, 1995; Des Jardins, 1990; Dettenmeier, 1996; Fishman, 1998; George, Light, Matthay, & Matthay, 1995; Traver, Mitchell, & Flodquist-Priestly, 1991), and emergency nursing, medical, and prehospital provider texts (Bledsoe, Porter, & Shade, 1994; Callaham, 1991; Caroline, 1991; Grant, Murray, & Bergeron, 1994; Hafen, Karren, & Mistovich, 1996; Holleran, 1994; Howell, 1998; Newberry, 1998; Rosen, 1998) found no uniformity in discussions of RD.

Because this proposal is concerned with persons cared for by OOH providers, OOH texts were studied in detail to determine the ways in which providers are being taught to evaluate patients with respiratory complaints. Four texts were for emergency medical technician (EMT) training (Caroline, 1991; Crosby & Lewallen, 1997; Grant et al., 1994; Hafen et al., 1996), two were for paramedics (Bledsoe et al., 1994; Caroline, 1995), and one was for nurses (Holleran, 1994). Although all texts used the phrase, "respiratory distress," only one text (Grant et al.) actually defined respiratory distress as "a combination of signs and symptoms that indicates a patient has a problem affecting the respiratory system...[it is] a series of indications that there is a problem" (p. 429). The phrase "respiratory distress" was used interchangeably in the prehospital texts with phrases such as "respiratory difficulty," "breathing problems," and "respiratory problems." Additionally, there was no uniformity among the texts regarding levels of respiratory distress. Descriptors of levels of respiratory distress or function included, "marked," "obvious," "adequate," "inadequate," "compensated," and uncompensated,"

in addition to "mild," "moderate," and "severe." No other text referred to the definition provided by Grant, nor was there any discussion of the need to establish an accepted definition or uniform assessment parameters.

This lack of consensus is illustrated in two recent publications reviewing the "state of the art" of OOH care for patients with asthma (Delbridge, Domeier, & Key, 2003) and acute pulmonary edema, the most common clinical manifestation of congestive heart failure (CHF) (Mosesso et al., 2003). As has been described, the respiratory distress of older patients with these pathologies is often so similar that they cannot be reliably differentiated in the OOH setting, yet these authors present somewhat different perspectives on how to assess the severity of that distress. Delbridge, Domeier, & Key suggest a variety of assessment parameters including the patient's general appearance, mental status, speech-to-breathing ratio, peak expiratory flow rate (PEFR), oxygen saturation levels, vital signs, and the providers' clinical judgments of severity. Mossesso and his colleagues suggest more general considerations such as hypertension, acute respiratory distress (undefined in the article), hypoxemia, tachypnea, rales, or wheezing. None of these authors suggests concrete guidelines for obtaining or interpreting the findings for any of the assessment parameters, nor did they offer standards for determining the severity of the respiratory distress.

Dyspnea and respiratory distress

Dyspnea is a well-recognized symptom of respiratory distress, and much time and effort has been spent trying to define it, measure it, and relate it to physiologic measures

and clinical outcomes. Although most people now consider dyspnea to be a symptom of respiratory problems, its use has varied both historically and currently (Ailani et al., 1999; Parshall, 1999). According to Parshall , the term "dyspnea" originated in antiquity and originally was considered a medical diagnosis. It has evolved throughout the centuries until the present when it is used, generally, to describe a person's subjective sense of the inability to breathe adequately. However, Alaini and others (1999) continue to use it to describe the general state of respiratory distress, including both subjective symptoms and measurable signs of that distress.

Although dyspnea is an important symptom, there is no clear correlation between the intensity of feelings of dyspnea and physiologic indicators of condition severity (Cook & Meek, 2001; Parshall et al., 2001; Wilson & Jones, 1989). The language of dyspnea, whether described by patients or providers, is no more precise than that of RD in general (Mahler, Jones, & Guyatt, 1998; Moy, Lantin, Harver, & Schwartzstein, 1998). As the patient's subjective sense of an inability to breathe adequately, dyspnea is one important indication of respiratory distress, but is inadequate, by itself, to define that distress.

Measuring respiratory distress

There are no valid, reliable measures of respiratory distress currently in widespread use in OOH or even emergency department settings. There are, however, several instruments available to evaluate dyspnea in patients (Ailani et al., 1999; Bestall et al., 1999; Hajiro et al., 1998; Karras et al., 2000; Kendrick, 2000; Mahler et al., 1998;

Moy et al., 1998; Wilson & Jones, 1989), and several more that have been developed for evaluating patients with croup (laryngotracheobronchitis) and asthma (Eisner et al., 1998; Jacobs et al., 1994; Janson-Bjerklie, Ferketich, Benner, & Becker, 1992; National Heart, Lung, & Blood Institute (NHLBI) 1997; Parkin, Macarthur, Saunders, Diamond, & Winders, 1996; Schuh, Johnson, Stephens, Callahan, & Canny, 1997; Smith, Baty, & Hodge, 2002; Steele et al., 1997; Taussig, Castro, Beaudry, Fox, & Bureau, 1975). In addition, one team attempted to develop an index to differentiate RD caused by pulmonary problems from that caused by cardiac problems (Ailani et al., 1999). All of these were reviewed critically (see below, Chapter 2, Literature Review), but none was suitable to be used in the OOH setting.

Summary

After thirty years of organized delivery of emergency medical services in the United States, there is general agreement that little is known about how ALS interventions affect either patient or system outcomes. In the case of respiratory distress, which has been identified as a priority condition for EMS outcomes research, there is not even an accepted definition of respiratory distress (RD), and there are no standards for assessing OOH patients, or for identifying patients in mild, moderate, or severe distress. In addition, at least one study has suggested that unless an older patient is in severe respiratory distress, erroneous treatment with ALS interventions could result in increased mortality rates. This proposed project, the development of a clinical decision rule to predict severe distress in older OOH patients in RD, addresses several of these

deficiencies. It begins with a definition of respiratory distress, proposes a standard assessment of OOH patients, and would establish the criteria for determining severe RD in older patients. Although this study represents only the very first steps in developing the Index of Respiratory Distress, it is critical to addressing issues in caring for older patients and to conducting meaningful outcomes research in this area.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter contains a critical review of the literature in three content areas: outcomes of OOH care; defining and measuring respiratory distress, especially as it relates to older OOH patients; and development and testing of clinical decision rules.

Outcomes of Out-of-Hospital Care

"Outcomes information...holds the promise of bridging the gap between what is done and what the intervention actually accomplishes," (White, 1997, p. vii). Research on clinical outcomes of OOH care is scant and what there is has failed to provide compelling evidence of the effectiveness of most ALS interventions

Brief History of EMS in the United States

It may be difficult for many persons to remember the way OOH care was delivered prior to the establishment of formal EMS systems in this country. As Howard (2000) pointed out, in the 1950's the only physical requirement for a vehicle to be called an ambulance was that someone could lie down in it, and of the 12,000 ambulances in service in the U.S. at that time, approximately 50% were owned and operated by morticians. Ambulance personnel were largely untrained, and what training did exist was haphazard and without standards. There were no communications between ambulances and receiving emergency departments, and the emergency departments themselves were unorganized and inadequately staffed.

The Emergency Medical Services Act of 1973 provided the bulk of the funding that enabled the growth of organized EMS systems throughout the country. These federal funds were augmented by sizable private foundation funding through the Robert Wood Johnson Foundation (NHTSA, 1996; Rose, 1980). In order to obtain funding, plans for EMS systems had to include 15 separate components: manpower, training, communications, transportation, facilities, critical care units, public safety agencies, consumer participation, access to care, patient transfer, coordinated patient record keeping, public information and education, review and evaluation, disaster plan, and mutual aid. While direct federal funding was discontinued in the early 1980s in favor of "block grants," EMS systems have continued to grow in this country, finding their funding through a variety of strategies.[3]

Emergency medical services systems continued to grow and change during the last 20 years of the 20th century, somewhat in parallel to the growth of emergency medicine and nursing. Present-day EMS systems include 14 recognized attributes (NHTSA, 1996) very different in their scope from the 15 components mandated as part of the original EMS legislation (see Table 1, below). These changes reflect a combination of cultural changes both within health care and beyond, the growth of knowledge in delivery of OOH and emergency care, changes in funding priorities, and changes in language.

[3] These options include general taxation, special assessments, direct fees for service, and combinations of these options.

Table 1

Comparison of Original Components of EMS with Present Attributes Abstracted from *EMS Agenda for the Future* [a]

Original Components	Present Attributes
Facilities	Integration of health services
Manpower	Human resources
Training	Education systems
Public information and education	Public education
Access to care	Public access
Communications	Communication systems
Critical care units	Clinical care
Coordinated patient record keeping	Information systems
Review and evaluation	Evaluation
Transportation	EMS research
Public safety agencies	Legislation and regulation
Patient transfer	System finance
Disaster plan	Medical direction
Consumer participation	Prevention
Mutual aid	

[a] Where possible, original components are paired with their current, similar attributes.

As Table 1 indicates, evaluation has been part of EMS systems since the original enabling legislation, and EMS research is a present attribute. Both evaluation and research have proven to be enormously challenging to EMS providers, administrators, and planners, and one of the greatest challenges in these areas is to define and measure appropriate outcomes, both at the individual patient level, and at the systems level.

Brief History of EMS Research in the United States

As was mentioned in Chapter 1, early EMS investigators recognized serious challenges to evaluating EMS care. Prime among those challenges has been, and

continues to be a lack of reliable outcome or performance measures. Spaite and colleagues (Spaite et al., 1995) echoed Gibson's (1973) earlier work when they cited a lack of system-based research as a continuing problem. They also identified problems inherent in the singular focus on the component-based medical model as leading to erroneous conclusions in evaluating OOH care. Nonetheless, investigators have pursued various solutions to the problem of defining and measuring outcomes of OOH interventions.

Defining and Measuring Outcomes of Out-of-Hospital Care

The two most widespread approaches to measuring outcomes of OOH care involve a) assessing various time intervals from EMS activation to patient arrival at the emergency department, and b) describing survival for victims of OOH cardiac arrest (CA). There is a good reason for this: these have been the easiest outcome measures to obtain. For this reason primarily, these approaches have been used to evaluate the effectiveness of treatments and the performance of EMS systems. However, these are far less than ideal measures of either system performance or the effectiveness of ALS interventions.

Use of response times to evaluate out-of-hospital care

From the time a call is made to a public safety answering point (PSAP) until the patient is delivered to an emergency department, the times for all events are entered in various places, including the PSAP and dispatch centers, the records of the OOH care

providers, and the emergency department records. However, there are problems associated with using even these meticulously documented times.

The first problem is that there are no good data relating response times to any patient outcomes other that survival from OOH CA (Hoekstra et al., 1993; Nicholl & Willoughby, 1998; Pepe, Abramson, & Brown, 1994; Spaite, Valenzuela, Meislin, Criss, & Hinsberg, 1993). Further, there is a great deal of evidence to indicate that measuring response times reliably in the OOH setting is difficult, if not impossible, despite widespread automatic/electronic monitoring (Cordell, Olinger, Kozak, & Nyhuis, 1994; Mosesso, 1993; Ornato et al., 1998; Spaite et al., 1995). Finally, except for victims of OOH CA, there is little reason to believe that differences in response time of 1-5 minutes (the increments usually cited as having an influence on survival of OOH CA) would make any difference in outcomes for patients with other problems (Callaham, 1997; Maio et al., 1999; Nichol et al., 1996).

Response time intervals may be of some use in evaluating EMS system function (Narad & Driesbock, 1999); here they would provide communities with an assessment of the level of service being provided. Although the problems inherent in measuring time intervals in OOH settings would still apply, using them to evaluate levels of service is a more straightforward and defensible application than using them as a clinical outcome measure.

Survival as an outcome of out-of-hospital care

Again, despite the ease of obtaining survival information, measuring survival as an outcome presents other challenges, some of which were recognized early in the history of OOH research and evaluation. As Rose (1980) identified, patients may either live or die regardless of the OOH interventions they receive. Further, survival may mean survival until arrival at the emergency department, survival until hospital discharge, or survival at various intervals post-hospitalization. Moreover, should investigators be concerned simply with survival or should they be concerned with the level of function the patient returned to after the incident? All of these variations appear in studies of OOH care, and efforts to standardize reporting, such as the use of the Utsein Criteria for research into survival from OOH cardiac arrest have not met with much success (Cummins, 1999; Swor, 1999).

Another significant problem with using survival as the primary outcome measure of ALS effectiveness is that the vast majority of ALS patients survive; OOH mortality is a rare event. Developing outcome measures for the vast majority of OOH patients is significantly more challenging than simply using indices of survival (Maio et al., 1999) but it is imperative if the effectiveness of OOH ALS interventions is to be evaluated meaningfully.

Recent Developments and Future Directions

Current knowledge of clinical outcomes of out-of-hospital research

In 1997, Callaham reviewed 5842 articles concerning OOH care and clinical outcomes that appeared in the MEDLINE database from January 1, 1985-September 1, 1997. That search yielded 170 clinical trials of any kind; with further review and a hand search of randomized trials, that number was reduced to 54 randomized clinical trials (RCTs) of interventions administered by non-physician OOH providers. Brice, and colleagues (2000) reviewed 285 articles of similar content that appeared in MEDLINE from 1985-1994, and they found 43 randomized trials and 10 non-randomized trials. Both of these studies arrived at similar conclusions: a) there is, in fact, very little known about the efficacy of most ALS interventions; b) there is a lack of well-designed studies; and c) outcomes other than mortality are rarely studied. Although both papers enumerate plausible reasons for these deficiencies, the problems remain unresolved.

The one area in which there is consensus that ALS interventions may make a difference is in survival from OOH CA (Callaham & Madsen, 1996; Maio et al., 1999; Nichol et al., 1996; Pepe, 1993; Pepe et al., 1994; Spaite et al., 1997). Interestingly, it appears that these results are due mostly to the immediate initiation of cardiopulmonary resuscitation – a Basic Life Support Skill – and early defibrillation – an ALS intervention that is rapidly becoming a BLS skill due to the development of fully automated defibrillators (Stiell, et al., 2003). Further, OOH CA represents only about 2% of most EMS activations (Callaham, 1997; Maio et al., 1999; Spaite et al., 1995), which means

that for the other 98%, we know even less about the effects of ALS intervention on patient outcomes.

Among the other 98% of OOH patients – those who did not suffer an OOH CA - are trauma patients. Given that EMS systems arose, in part, in response to a perceived need to improve the OOH care of trauma patients (Howard, 2000), it is interesting to note that what improvements have been made have, again, resulted largely from BLS interventions (Eckstein, Chan, Schneir, & Palmer, 2000; Liberman, Mulder, & Sampalis, 2000; Spaite, Criss, Valenzuela, & Meislin, 1998). As Liberman, Mulder, & Sampalis concluded, "The aggregated data in the literature have failed to demonstrate a benefit for on-site ALS provided to trauma patients..." (p. 584).

What is known about the outcomes of OOH care is limited almost exclusively to the survival rate of cardiac arrest and trauma patients. This leaves a great deal to be done to determine other outcomes of other OOH interventions for other kinds of patients.

EMSOP, OPALS and recent OOH outcomes research

The interest in health care outcomes research that emerged in the mid-nineteen nineties affected OOH practice much as it did all other aspects of the health care industry. Although Callaham's (1997) article describing the paucity of high-quality research in OOH care was one of the earliest and most provocative discussions addressing this problem, others had recognized the need to expand the knowledge base in this area, and two large projects were initiated in the late 1990s to address these needs. The Emergency

Medical Services Outcomes Project (EMSOP) was undertaken in the United States, and in Canada, the Ontario Prehospital Advanced Life Support (OPALS) project was begun.

The National Highway Traffic Safety Administration, the lead agency in the U.S. federal government for EMS, convened a meeting in 1994 to address these problems. The EMSOP group described above was a major result of this meeting. To reiterate, the three goals of the EMSOP project were to set priorities for outcomes research, develop appropriate risk adjustment measures, and develop sensitive outcome measures for the highest priority conditions.

The process for setting priorities for OOH research was described above in Chapter 1. Table 2, below, presents the top 10 conditions that were identified by the EMSOP group as priority conditions for OOH research. These conditions were assessed as prevalent in EMS care, amenable to OOH interventions, and likely to be positively affected by such care (Maio et al., 1999)

Table 2

Top Priority Conditions Identified by the EMSOP Group for Out-of-Hospital Research

Priority Number	Adults	Children
1	Minor trauma	Minor trauma
2	Respiratory distress	Major trauma
3	Chest pain	Respiratory distress
4	Major trauma	Airway obstruction
5	Cardiac arrest	Respiratory arrest
6	Shock	Cardiac arrest
7	Diabetes complications	Seizure
8	Allergic reaction	Shock
9	Environmental exposure	Allergic reaction
10	Stroke/cerebrovascular accident	Environmental exposure

After identifying the priority conditions for OOH outcomes research, the EMSOP group then developed a "methdological foundation" to guide future work in this area (Spaite et al., 2001, p. 658). This is described more fully below in the section describing the conceptual framework of this study.

In their third report (Garrison et al., 2002), the EMSOP group described the concept of risk adjustment measures (RAMs), discussed the relevance of such measures in OOH research, and offered some examples of RAMs for standard use in OOH research. They identified what they called core RAMs that should be assessed for all conditions, and specific RAMs, those collected for specific conditions. Examples of core RAMs they identified included such items as age, sex, race or ethnicity, and vital signs before and after interventions. By way of example, they suggested the following RAMs

be used for all patients with complaints of respiratory distress: peak expiratory flow rate, pulse oximetry, and visual analogue scale measure of dyspnea.

The Ontario Prehospital Advanced Life Support (OPALS) study is an on-going, large, multi-site study looking at outcomes – including cost-effectiveness – of OOH care in over 40,000 critically ill and injured patients in an urban EMS system (Stiell, DeMaio, Nesbitt, et al. 2003; Stiell, Wells, et al., 2003;Stiell, Wells, DeMaio et al., 1999; Stiell, Wells, Field et al., 1999). Their work initially focused on outcomes of patients suffering from OOH CA, (Stiell, Wells, DeMaio et al., 1999; Stiell, Wells Field et al., 1999) but has recently included a variety of outcomes for patients in acute respiratory distress (Nichol et al., 2003; Stiell, Wells et al., 2003). These include a) clinical outcomes such as OOH provider judgment of improvement associated with treatment, intubation and mortality in the ED, radiologic evidence of aspiration, hospital length of stay, and best clinical assessment at discharge (Stiell et al., 2002); b) survival (Stiell, DeMaio, Nesbitt, et al., 2003); and c) the functional outcome of health-related quality of life (Nichol et al., 2003).

The work of the OPALS group is particularly interesting, and many of these individuals have also been active in developing and evaluating clinical decision rules (described below). It is clear that this group also struggles to identify appropriate outcomes against which to evaluate OOH interventions. There are, for example, problems with many of the clinical outcomes they proposed for patients in acute RD. First, the judgment of EMS providers has no inherent meaning and its reliability is questionable.

Second, length of hospital stay and clinical condition at discharge are both temporally removed from OOH interventions to an extent that challenges causal inference. Third, health-related quality of life is truly remote from OOH interventions for acute respiratory distress, and the "stability" of this measure as reported in their study most probably reflects the fact that overall quality of life is little impacted by one acute episode of respiratory distress.

For all of the positive that is emerging from the EMSOP and OPALS projects, there remain two generic problems that will continue to confront OOH outcomes investigators, and those are the difficult work environment of OOH care, and ethical concerns, particularly around issues of informed consent. It may turn out that it will not be feasible, either ethically or practically, to conduct RCTs to evaluate currently accepted OOH ALS interventions. In many cases these practices are now considered "community standards," and withholding them may not be acceptable even if their effects on outcomes are uncertain. In those cases, well-designed case-control studies may have to suffice. Without RCTs, it is impossible to reliably assess the impact of new interventions on any of the six outcomes identified in the EMSOP study (Callaham, 1997; Maio et al., 1999). This has resulted in the unfortunate situation where OOH ALS interventions are introduced as changes in protocols (which are not subject to the scrutiny of institutional review boards), and then are evaluated retrospectively, if at all, as part of the system's ongoing evaluation program.

The potential contribution of a well-designed clinical trial of OOH ALS interventions was recently demonstrated by Gausche and her colleagues (Gausche et al., 2000) in a RCT comparing survival rates and neurological outcomes of pediatric patients who were intubated in the field with those who received bag-valve-mask assisted ventilation.[4] In this study, paramedics received extensive training in both pediatric endotracheal intubation (ETI), and in the correct management of bag-valve-mask (BVM) assisted ventilation. The study took place over three years, and 830 consecutive patients aged 12 or younger were enrolled; assignment to treatment groups was made by day, with 410 enrolled on odd days (BVM) and 420 enrolled on even days (ETI). Survival and neurological outcomes were evaluated retrospectively using well-developed study protocols. The results of this study demonstrated that ETI did not, in fact, improve outcome in the urban setting in which the study was conducted; in addition, the BVM group had significantly fewer complications than did the ETI group. As a result of this study, the research team recommended against using ETI with pediatric patients in urban settings with short response times. Although not a part of the EMSOP project, this study illustrates nicely how well-designed OOH outcome studies can be used to assess both the benefits and risks of OOH ALS interventions *prior to* implementation.

---

[4] The study received institutional review board or equivalent approval from all 115 receiving facilities in the study, as well as the sponsoring agency, with a waiver of consent for the patients enrolled.

Assessing Respiratory Distress in Out-of-Hospital Patients

Defining respiratory distress

As was discussed above in Chapter 1, there is no accepted definition of RD, nor are there any standards for assessing that distress or assigning a level of severity. For the purposes of this study, RD is defined as, *that combination of signs and symptoms that indicates a person's physiological and behavioral responses to actual or perceived problems with ventilation, oxygenation, or respiratory effort.* The domains included in this definition, i.e., ventilation, oxygenation, and respiratory effort, were derived from a review of selected medical and nursing literature.

Medical pulmonary texts identified arterialization of venous blood, mechanics of respiration, and contraction of muscles of inspiration as the components of respiration (Dantzker et al., 1995; Levitsky, 1999). General and pulmonary nursing texts identified gas exchange, inhalation and exhalation, and the need for adequate strength in muscles of respiration (Dettenmeier, 1996); ventilation, alveolar diffusion, transport of gases, and control of ventilation (Wilson & Thompson, 1990); and ventilation, perfusion, and control of breathing (Thompson et al., 1997). Finally, in selected prehospital texts, important components of respiration were identified as inhalation, exhalation, and oxygenation (Crosby & Lewallen, 1997; Hafen et al., 1996).

The domain *ventilation* was identified as best description of movement of air into and out of the lungs; it comprises both inhalation and exhalation. Although gas exchange covers the movement of both oxygen and carbon dioxide, the domain *oxygenation* is the

critical component of gas exchange in patients with acute respiratory problems. Finally, *respiratory effort* is the domain that effectively identifies the work of breathing that patients with acute respiratory distress are either able or unable to accomplish.

Measuring respiratory distress

As was mentioned in Chapter 1, many attempts have been made to evaluate the severity of RD in selected patient populations, primarily in children with asthma and croup. These were all reviewed to assess their suitability for use in the OOH setting.

Dyspnea scoring instruments that allow patients to self-rate the intensity of their symptoms have been developed and tested. The two most common instruments are visual analogue scales, and the modified Borg scale (Hajiro et al., 1998; Kendrick, 2000; Mahler et al., 1998; Wilson & Jones, 1989). Although these are widely accepted for use in assessing dyspnea, dyspnea is only one manifestation of respiratory distress. In addition, these scales were developed for use primarily with patients experiencing chronic problems, and their use with patients in acute respiratory distress has not been widely evaluated.

One study, however (Moy et al., 1998), has reported on the use of a dyspnea rating scale with emergency department patients experiencing acute respiratory distress from an asthma attack. That study reported that dyspnea was an important component of respiratory distress and that evaluating it may be critical to making appropriate clinical judgments. Kendrick (2000) also proposed using a modified Borg scale to assess dyspnea

in asthma patients in the ED, but her report was purely anecdotal and did not include any measures of reliability or predictive validity associated with it.

Ailani and colleagues (1999) reported the development of a "dyspnea differentiation index," (DDI) designed to differentiate dyspnea of cardiac origin from that of pulmonary origin. This research team measured what they called "dyspnea" using peak expiratory flow rates (PEFR) and partial pressure of oxygen in arterial blood (PaO_2) to derive a measure that they hoped would discriminate between patients in congestive heart failure (CHF) and those with chronic obstructive pulmonary disease (COPD). Their report contains detailed information on the development and testing of the instrument; however, their efforts met with only modest success. The DDI was only marginally better at differentiating the causes of dyspnea than the PEFR itself, and neither was significantly better than the unaided judgment of the initial examining physician. It was not clear from this report that the DDI added any useful information to the decision-making process in the emergency department, and its reliance on the PaO_2 makes it especially unsuitable for use in OOH care.

It is interesting to note Ailani and colleagues'(1999) use of PEFR and PaO_2 as measures of dyspnea. As has already been discussed, dyspnea is considered a self-report of patients' experiences of breathlessness, and it is usually assessed using a visual analogue scale or other standardized self-reporting instrument such as the modified Borg scale. Although these investigators claim to be measuring dyspnea, they were, in fact, attempting to assess patients' ventilation and oxygenation status using standard

physiological measures. In some ways, this may represent a first step in attempting to differentiate the kind and amount of *respiratory distress* in patients with CHF vs. those with COPD.

The instruments described by Janson-Bjerklie, et al. (1992)and Eisner, et al.(1998), while they measured parameters in addition to dyspnea, were also inappropriate for use with OOH patients. Rather than focusing on acute RD, these measures were developed to evaluate disease severity in patients with chronic asthma. As such, they included information such as numbers of previous hospitalizations, patients' estimations of their own risks of death from their disease, etc. In addition to focusing on long term signs and symptoms, these instruments are complex, and would be impossible to complete in most OOH settings.

The remaining instruments described in the articles by Schuh, Johnson, Stephens, Callahan, & Canny (1997), Parkin, MacArthur, Saunders,Diamon, & Winders (1996), Smith (2002), Steele, et al. (1997), Jacobs, et al. (1994), and Taussig, Castro, Beaudry, Fox, & Bureau (1975), as well as the National Heart, Lung, and Blood Institute (NHLBI) classification scheme (1997) may all be helpful in selecting appropriate items for any evaluation of RD in OOH patients, however, none is adequate to be used as is. Except for the NHLBI scheme, these instruments were developed in emergency departments to evaluate the severity of asthma or croup in children of varying ages. They were designed to serve one or more of the following purposes: a) to assess patients' responses to therapy; b) to provide guidelines for admitting patients to the hospital and/or the critical

care unit; or c) to test a particular component of the evaluation as being a single best measure of severity. Few of these instruments were subjected to any formal reliability or validity testing, and none was used on older adult patients, or in any out-of-hospital setting. A more detailed discussion of these instruments can be found in Appendix 1.

The instruments and decision rules evaluated here all provide some information that will be useful in developing the Index of Respiratory Distress. However, none is appropriate for use with older patients in the OOH setting.

Clinical Decision Rules

General Considerations

Clinical decision rules, as described above, are tools designed to assist providers in clinical decision-making. They consist of three or more predictor variables derived from the patient's history, physical examination, or laboratory findings, and an outcome variable that is either the probability of a diagnostic outcome or a prescribed treatment. (Laupacis et al., 1997; McGinn et al., 2000; Stiell & Wells, 1999; Wasson et al., 1985). All of these authors, many of whom worked together to develop these criteria, argue that unless CDRs are developed and tested in a rigorous manner, they will fail to give reliable results and would thus add nothing of value to clinical practice. Taken together, the works of these authors provide guidelines for developing and evaluating CDRs.

Criteria for Developing and Validating Clinical Decision Rules

Stiell & Wells (1999) and McGinn, et al. (2000) present six criteria for assuring the adequacy of a CDR, and Stiell & Wells explicate each further with more detailed

"standards" by which to assess it. The first criterion is a demonstrated need for such a rule. Need may be demonstrated by such things as the clinical prevalence of the problem, wide variation in practice, clinical accuracy of providers, or the resources required for a particular test or procedure. In OOH practice, for example, clinical accuracy is a major issue, where providers with varying levels of education and experience work in adverse conditions with few reliable diagnostic aids.

The second criterion proposed was that the rule be derived according to some methodological standard (Stiell & Wells, 1999). Standards included specified guidelines for defining both the predictor and outcome variables, suggested reliability standards for the predictor variables, issues associated with sample selection and size, and various statistical techniques appropriate for developing and testing the rule. The authors also insisted that the internal validity of the rule be demonstrated by its consistency, accuracy, and clinical sensibility.

The third criterion, that the rule be prospectively validated and refined, is critically important in establishing the external validity of the CDR. The authors pointed out that the failure to determine a CDR's accuracy in a new patient population has resulted in the failure of these rules to perform well when tried in settings other than where they were developed (Altman & Royston, 2000; Stiell & Wells, 1999). Standards associated with this rule include caveats regarding sample size and selection, establishing a "gold standard" outcome against which the CDR outcome would be evaluated,

assessing the reliability and accuracy of the rule, and assessing the potential effects of the rule, including any potential cost savings associated with its use.

Following prospective validation, the next criterion is limited successful implementation into clinical practice. Stiell & Wells (1999) advocate the use of RCTs in a variety of settings to test the rule's effectiveness. The rule needs to be evaluated, again, for its accuracy, and most importantly, at this stage, the rule should be assessed for acceptability among clinicians.

In the final criterion, Stiell & Wells (1999) asked, "How will the rule be disseminated and implemented?" (p. 445) In this criterion, they enumerated some of the attributes that have been demonstrated to increase the adoption of innovations in practice, including demonstrated improvement over old practice, ease of use, similarity to past practice, and the ease with which the rule may be "tried out," before being implemented. The six criteria are summarized in Table 3, below.

Table 3

Criteria for Evaluating and Developing Clinical Decision Rules[a]

1. Is there a need for the decision rule and how is that need demonstrated?
2. What were the methodological standards used to develop the rule?
3. Was the rule prospectively validated and refined?
4. How has the rule been successfully implemented into clinical practice?
5. Will the rule be cost-effective, and how was that evaluated?
6. How will the rule be disseminated and implemented?

[a](adapted from Stiell & Wells, 1999, p. 438)

Problems with Clinical Decision Rules Currently Used in OOH Practice

Identifying and measuring predictor variables

Trauma triage guidelines are examples of CDRs in widespread use in the OOH setting, and they provide a good example of the difficulties in identifying and measuring predictor variables in this environment. Trauma triage guidelines are designed to determine which patients are transported to trauma centers and which are cared for in other facilities. The predictor variables used in these instruments include physiological variables such as pulse rate, blood pressure, respiratory condition, and skin color. Some guidelines add anatomical predictor variables such as bleeding, injury region, and injury type (Cales, 1986), and some have added assessments of mechanisms of injury as well, including, for example, falls from a prescribed distance, or motor vehicle impacts at a

prescribed speed (Baxt, Berry, Epperson, & Scalzitti, 1989). Most include either all or part of the Glasgow Coma Scale (GCS) – a CDR in its own right – to evaluate the extent of any head injuries. Finally, other variables may be added with the goal of increasing the CDR's predictive validity. These include things like patients' ages or co-morbidity (Baxt et al., 1989; Cales, 1986; West, Murdock, Baldwin, & Whalen, 1986). Typically OOH providers assess patients according to these predictors and derive a score. That score is compared with a critical cut-off; the position of the patient's score relative to the cut-off determines which emergency department will receive the patient.

One of the most important components in any CDR is the consistency of the predictor variables. Laupacis and colleagues (1997) acknowledge special difficulties associated with clinical predictor variables, and the GCS, a key component in most trauma triage rules, provides an excellent example those difficulties.

Predictor variables used for the GCS include the patient's best motor response, best eye-opening response to verbal commands or painful stimuli, and best verbal response to commands. Each of these is rated on a scale of 1-4, 1-5, or 1-6 respectively, with a total possible score ranging from 3, indicating very poor neurological function, to 15, indicating no neurological deficits. Table 4, below, shows the GCS.

Table 4

Glasgow Coma Scale[a]

Activity	Best Response	Score
Eye Opening	Spontaneous	4
	To verbal stimuli	3
	To pain	2
	None	1
Verbal	Oriented	5
	Confused	4
	Inappropriate words	3
	Nonspecific/unintelligible sounds	2
	None	1
Motor	Follows commands	6
	Localizes pain	5
	Withdraws from pain	4
	Flexion to pain	3
	Extension to pain	2
	None	1
	Possible total score	3-15

[a](Adapted from Harrahill, 1996)

The problem with these predictor variables is that they are neither clearly defined nor able to be reliably measured. In their report on the Management and Prognosis of Severe Traumatic Head Injury (Brain Trauma Foundation (BTF), 2000), the BTF did an extensive, evidence based review of the GCS and its use in assessing traumatic head injury. Among its findings was that the reliability of the initial ratings, especially those done in the OOH environment, could not be established. There are several problems associated with the use of this scale, but one of the most important is the high inter-rater

variability in assessment of the predictor variables and the resulting variability in total scores.

Laupacis, Sekar, & Stiell (1997) acknowledge the difficulty in obtaining reliable results from clinical evaluations. These authors accept kappa scores of as low as 0.6 as evidence of acceptable inter-rater reliability in CDRs that rely on these kinds of predictor variables. Although neither Baxt et al. (1989), nor Cales (1986) directly addressed the problems related to the reliability of the providers' assessments of physiological variables in the trauma triage guidelines they evaluated, these problems are the basis of the lack of predictive validity that they and other investigators have identified (Maslanka, 1993; Senkowski & McKenney, 1999).

Identifying and measuring outcome variables

The problems associated with defining and measuring outcome variables in OOH practice are comparable to those of predictor variables. In their initial presentation of the GCS, Teasdale & Jennett (1974) offered their scale as a means of standardizing assessment of the neurologically impaired patient, and of defining "coma" more precisely. Other than accurately measuring the depth of coma (an outcome for which there was no "gold standard" against which to evaluate it), they did not have a particular "outcome" in mind. They did acknowledge that their scale might improve the ability to predict the ultimate outcome in patients with varying levels of coma, particularly after head injury, however, this was clearly not the intended purpose of the scale, nor did they do any formal predictive testing. Today, although it is frequently used to predict

outcomes of head-injured patients, the GCS fails to do so reliably; there are significant numbers of patients who recover almost completely from what appear to be significant traumatic injuries with associated low GCS scores (BTF, 2000).

Trauma triage guidelines were intended to identify critically injured trauma patients. The outcomes most often used to evaluate the success of these CDRs include mortality, need for emergency surgery, intensive care unit stay, overall hospital stay, and measures on indices of injury severity that themselves have been criticized on many levels (Baxt et al., 1989; Cales, 1986; Lett, Hanley, & Smith, 1995; Maslanka, 1993; Senkowski & McKenney, 1999). Each of these outcomes is influenced by a host of factors that may or may not be related to the trauma triage decision. There simply is no "gold standard" against which any of the trauma triage guidelines can be compared. In fact, one widely cited study (Emerman, Shade, & Kubincanek, 1991) indicated that certain trauma triage guidelines may be no better at determining mortality or need for emergency surgery than the unaided judgment of OOH providers.

In a recent contribution, Newgard, Lewis, & Jolly (2002) reported on their efforts to create a CDR that could be used to accurately predict severe injury in pediatric trauma patients. They used as their predictor variables only those measures currently available to OOH personnel, and their outcome variable was an Injury Severity Score (ISS) of less than 16. They used prospectively collected data from a national probability sample of patients involved in motor vehicle crashes (MVCs). Their sample of over 8000 patients was based on data collected from 1993-1999. Using the kinds of multivariate analyses

suggested as appropriate for developing CDRs, they derived a rule using three variables: the Glasgow Coma Scale score, intrusion of the vehicle into the passenger space of six inches or more, and lack of proper passenger restraint. They reported a sensitivity of 92% and a specificity of 73% with their rule.

Newgard, et al (2002) did not specifically address the reliability of their predictor variables, but it is interesting to note that they chose to use the initial GCS *from the receiving facility* in their analysis rather than that recorded by the OOH providers. In fact, none of the predictors they chose involved a clinical assessment by OOH providers. Rather, their predictors included such items as age, sex, and weight of the patient; primary point of impact on the vehicle and the extent of passenger space intrusion; and seat location and restraint use. Their chosen outcome variable, the ISS, has been used in other studies, but that score is derived after the episode of care is completed and may not be the best reflection injury severity at the time of the crash. They do suggest the need to validate their findings in subsequent prospective studies, and such studies will be important in evaluating the utility of this rule in actual practice.

Problems with both the predictor and outcome variables represent clear limitations with both the GCS and trauma triage guidelines as currently used in OOH practice. They also illustrate quite clearly some of the challenges to developing the Index of Respiratory Distress for use by OOH providers.

Conceptual Framework

Spaite and colleagues (2001) developed a conceptual model for OOH research that addresses several issues inherent in research in this setting. One of the most vexing problems in OOH outcomes research is the issue of causal inference. This refers to the fact that the outcomes of care most frequently assessed are things like survival, length of stay in the emergency department, admission to the hospital, length of stay in the hospital, and costs of care. With the exception of survival from OOH cardiac arrest, it is difficult to demonstrate that OOH care has any direct impact on any of these outcome measures. Another related challenge to OOH researchers involves isolating OOH interventions within the entire interaction between the patient and the health care system. The model described below addresses these issues by breaking down the EMS-patient interaction into smaller, more easily defined components.

In response to these challenges, the authors (Spaite et al., 2001) developed what they called the "Episode of Care Model" (p. 659), in which all aspects of a single EMS activation are contained within five components, known as "units of service": OOH care, emergency department care, emergency sub-specialty care, inpatient care, and follow-up care. Each episode of care is initiated by a precipitating event and each results in one or more short and long-term outcomes. Outcomes research directed at any episode of care, e.g., acute respiratory distress, would necessarily be focused on one or more units of service within that episode, and for each such unit of service, the authors suggest that

investigators must address risk adjustment measures, therapeutic interventions, and outcome measures. Figure 1, below, illustrates how this might look.

Figure 1

Example Using the Episode of Care Model for Acute Respiratory Distress

Unit of Service	**Examples of Risk Adjustment Measures**	**Examples of Therapeutic Interventions**	**Examples of Short- and Long-term Outcomes**
OOH care	Age, comorbidity, initial level of respiratory distress	Oxygen, albuterol, nitroglycerine, furosemide, morphine	Survival, level of respiratory distress at arrival to ED
ED care	OOH interventions, severity on arrival to ED	CPAP/BiPap[b], ETI[c], medications	Survival, ED length of stay, inpatient admission
ED subspecialty care	ED interventions	Antibiotic therapy, chest tube insertion	Survival, admission to ICU
In-patient care	All previous RAMs[a]	Continued therapy specific to condition	Survival, length of stay in ICU, length of stay in hospital
Follow-up	All previous RAMs	Post-hospital discharge interventions	Survival, functional capacity, cost of care

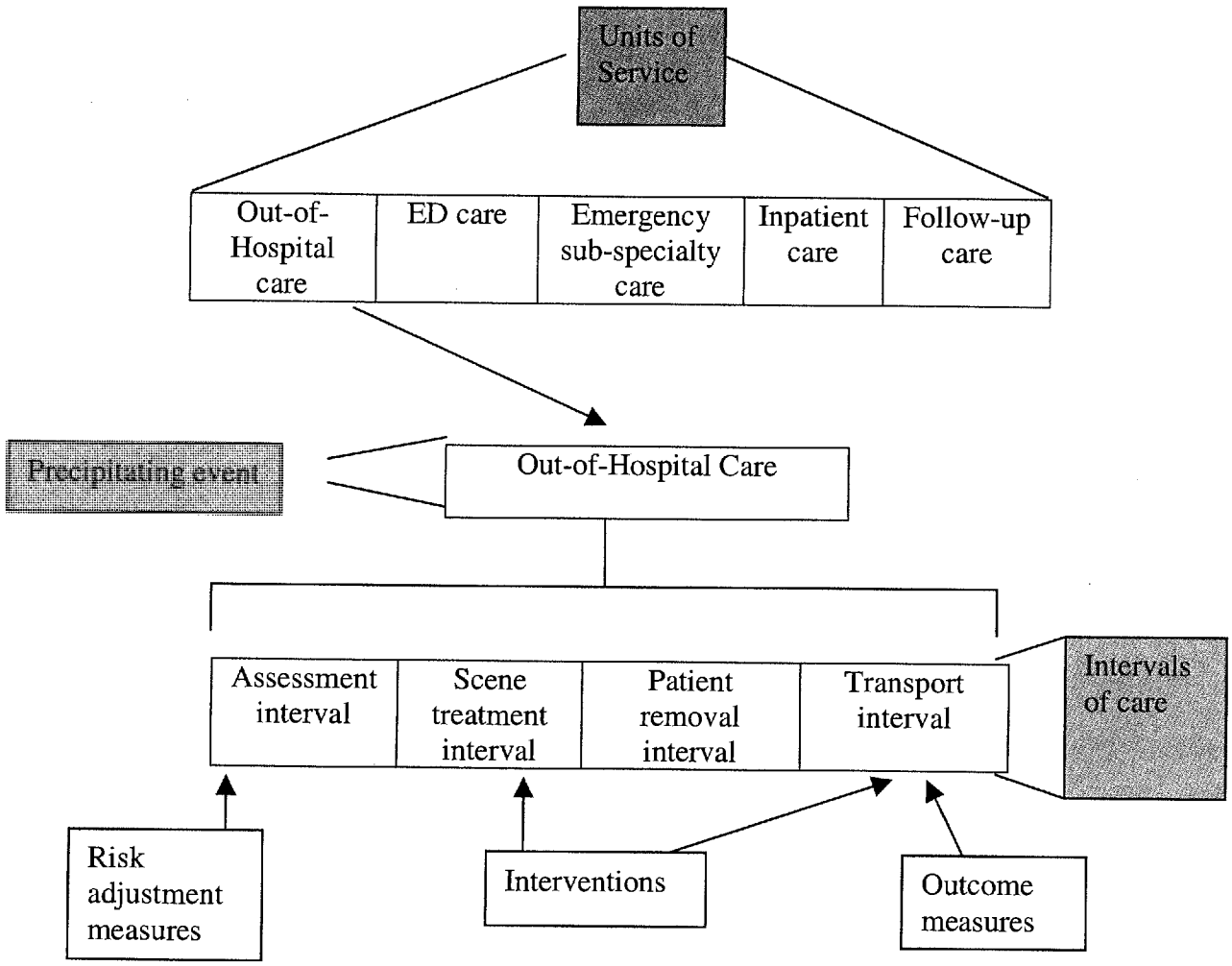
[a] RAM = risk adjustment measure
[b] CPAP=continuous positive airway pressure; BiPAP= Bi-level positive airway pressure
[c] ETI= Endotracheal intubation

Several things become apparent immediately in considering this model. First, survival – or conversely mortality - is an outcome that may be measured in all units of service, and may reflect outcomes of interventions at any level, or as previously mentioned, may be simply the unavoidable outcome of the underlying disease. Second, the "distance" from the OOH interventions to important outcomes such as functional capacity and cost of care make it incredibly difficult to design any research that could control all the intervening variables enough to describe the effects of OOH interventions on these important outcomes.

In this same article (Spaite et al., 2001), the authors described the OOH unit of service in even finer detail, including components of the assessment interval, the scene treatment interval, the patient removal interval, and the transport interval. They identified the assessment interval as appropriate for determining risk adjustment measures, the scene treatment and transport intervals as appropriate for interventions, and late in the transport interval as appropriate for determining the outcomes attributable to the OOH interventions. Figure 2, below, illustrates the more complete version of the model.

Figure 2

Conceptual Model for Out-of-Hospital Outcomes Research[a]



[a]From Spaite, et al. (2001)

According to this model, each EMS activation results in care that can be placed in one or more of the *units of service* pictured above; OOH care is the unit of interest here. The EMS activation occurs as a result of a *precipitating event*, pictured at the start of the expanded OOH unit of service. The OOH care unit is comprised of four *intervals of care*, each of which consists of a relatively distinct component of OOH care. As described above, research in the OOH care unit must include *risk adjustment measures* which are obtained during the assessment interval, *interventions*, which may be applied at the scene or during transport, and *outcome* measures, which are obtained as the OOH care unit of service is completed and the patient moves into the ED unit of service.

The Index of Respiratory Distress will fit into this model in both the assessment interval and the transport interval. In the assessment interval, OOH providers would use the IRD to assess the initial level of severity of distress their patient is experiencing. Theoretically, the results obtained using the IRD would help guide the interventions used in the scene and transport intervals. The IRD would be used again in the transport interval to determine any changes in the patient's level of respiratory distress, especially in response to the interventions applied.

Although the IRD is being developed primarily to assist in the clinical care of patients, it will also make some significant contributions to OOH research with older patients in respiratory distress. In some sense, the use of the IRD as part of the initial assessment could serve as a risk adjustment measure. Older patients activate the EMS system for respiratory problems at different levels of illness/distress. Studies designed to

assess the efficacy of various OOH interventions will have to account for these differences in a systematic way in order to determine if the observed outcomes are a result of the interventions or a function of the severity of the patients' illnesses. At the present time, there is no single, standard measure to account for these differences; the IRD could possibly fill this gap.

CHAPTER 3

DESIGN AND METHODS

This is a mixed-methods study, using qualitative methods to select items appropriate for the IRD, and quantitative methods for determining the best predictors variables. This chapter is presented in two sections: data collection and analysis. Within each section, the methods are presented in chronological order, beginning with methods used to select items for the IRD, continuing with the method used to reduce the items to a number suitable for evaluation using statistical methods, and ending with methods used for determining the best predictors of RD in older OOH patients. Oregon Health & Science University's (OHSU's) Institutional Review Board (IRB) approval for the study was obtained in two phases: first for the observations and interviews with OOH providers, and second for abstracting data from both the OOH and ED medical records. A discussion of human subjects considerations immediately precedes the section on Analysis.

Data Collection

Item Selection

Literature review

The textbooks and research literature cited in Chapter 2 were reviewed and descriptions of clinical indicators of respiratory distress were identified. This review included the reports of various CDRs used to assess patients with asthma and croup.

Review of out-of-hospital records

The investigator reviewed 28 de-identifed patient records from one of the local participating EMS agencies in order to determine the data that providers were recording as they assessed older patients in respiratory distress. This review provided information about how OOH providers were actually assessing patients and recording their findings. The agency provided complete records, including both the narrative section in which the OOH providers wrote in longhand, and the computer-ready "scantron" form, in which providers blackened areas corresponding to certain assessment parameters. Appendix 2 contains a copy of both sides of one such form; all patient and provider identifying information has been removed.

Observations and interviews with out-of-hospital providers

It has been the investigator's experience that there is much that contributes to decision-making among health care providers that is not documented fully in medical records. The literature describes how assessments should be performed, how providers should interpret their findings, and how those findings should lead to treatment decisions. However, that may not be what actually happens in practice. To ensure that the IRD would have relevance to OOH providers, the investigator sought to ride along with them and observe their actual practice.

Per the process approved by the OHSU IRB, the investigator contacted key personnel at three EMS agencies in the Portland metropolitan area. This was a convenience sample of agencies, representing both city and suburban regions, as well as

first responder agencies and transport agencies. All three agencies agreed to participate. Two of the agencies were fire departments that provided first responder ALS services but did not transport patients, and one agency was a private provider that transported ALS patients to area hospitals. The investigator met with managers and training personnel at each agency and described the study in detail. She requested these key personnel identify providers in their agencies with several years' experience who might be interested in participating. Key personnel made initial contact with the providers, described the study, and sought permission for the investigator to contact them. Only after the providers indicated an interest in participating in the study were the ride-alongs scheduled. In some agencies, the investigator was able to contact providers at work ahead of time to confirm their willingness to participate; in other agencies, work routines did not permit that initial contact. In all cases, each time the investigator arrived for a ride, she explained the study in detail, reviewed the risks and benefits to the participant, emphasized issues of confidentiality, and stressed the voluntary nature of their participation. After answering any questions the participants had, she obtained written informed consent from them. A copy of the consent form appears in Appendix 3.

When the OOH providers were caring for older patients in RD, the investigator took notes on providers' behaviors, focusing on the questions they asked patients, the assessments they performed, and the care they provided. After care of the patient was transferred to emergency department personnel, the providers were debriefed to help clarify what the investigator saw and heard. The focus of the debriefings was on what the

providers saw, why they did what they did, what conclusions they drew from their assessments, and their thinking processes throughout the encounter. The investigator took notes during the debriefings. Appendix 4 provides a sample of the guidelines the investigator used when observing OOH providers.

In addition to de-briefing the providers after patient encounters, the investigator conducted informal interviews with providers, asking them to describe how they could tell when older patients in RD were "really sick." Specifically, what parameters did they assess, and how did they interpret the results of those assessments? The goal of these interviews was to determine what signs and symptoms the providers said they relied on as key indicators of severe distress in patients with respiratory complaints. The investigator was careful not to make any suggestions or refer to statements by other providers; rather, she prompted providers to tell her what ***they*** used to determine for themselves whether or not patients were in severe distress. The investigator made notes of the responses during these interviews; no tape recording was done. Appendix 5 contains the guidelines the investigator used for these informal interviews.

After compiling a list of clinical indicators from the literature, EMS records, and OOH providers, it was necessary to condense the list to allow for reasonable sample sizes for statistical testing.

Item Reduction

Item reduction was accomplished using the opinions of selected experts in OOH practice (Streiner & Norman, 1995). A group of eight providers with extensive

experience in OOH care was identified by the investigator; three of these experts were physicians, three were OOH providers, and two were nurses. Five of these experts were from the Portland metropolitan region; the other three were from other regions of the country, including northern CA, Michigan, and Pennsylvania.

The investigator contacted each expert either in person or by phone. She explained the purpose of her study and what their role might be. Because no identifying information was being sought from the experts, and because they were not considered participants in the study, the OHSU IRB agreed that no consent was required. All the experts agreed to participate.

A Delphi technique (Polit & Hungler, 1999) was used to elicit the experts' opinions on the best clinical predictors of severe respiratory distress from among those presented in each mailing. This technique uses an iterative process to reach consensus among a group of experts. The experts are surveyed on the topic, and the results are tabulated. The results are then distributed to the experts, who are questioned again based on the results of the first survey. The process continues until consensus is reached.

The first survey sent to the experts consisted of a packet containing the following: a) a cover letter reiterating the purpose of the study and their role in it, b) a sheet asking for information regarding their expertise in OOH care, c) three rating sheets, d) a final sheet asking for any additional suggestions from the expert, and e) a return envelope. A copy of this packet is contained in Appendix 6.

The experts were asked to rate each of the clinical indicators on three different dimensions. The first was the importance of the indicator as a measure of the severity of disease; the second was how likely they thought OOH providers were to be able to assess that indicator in older patients in a rapid, consistent manner in the OOH environment (either at the scene or en route); and the third was which indicators they thought OOH providers used routinely in assessing their older patients in respiratory distress. These dimensions were referred to as severity, reliability, and usability. The experts were asked to score each parameter in each dimension on a scale of 1-3, with 3 being the best in each category and 1 being the worst.

In order to complete the rankings and reduce the list to its final form, a second survey form was mailed to the experts (Appendix 7). In this second survey, the experts were asked to rank order the remaining indicators with the rank of 1 indicating the best potential predictor of severe distress and 15 indicating the worst. The investigator consciously omitted any suggestions for how the experts were to determine the "best" potential predictors; she was looking for their expertise based on whatever criteria they used. Experts were also asked to identify any items they thought didn't belong on the list (and to indicate why they thought it didn't) and add any they thought were critical but didn't make it on to the final list.

Item Scaling

The scaling of indicators identified by the experts and retained for the final selection was kept simple, and in line with other CDRs that OOH providers already use.

The range of possible scores for each item was from 2 indicating normal findings (or no deviation from the patient's baseline), 1 indicating mild alteration from normal, and 0 indicating significant alteration from normal findings. This is similar to the Glasgow Coma Scale, or many trauma triage rules in which lower scores for each item (and for the total) indicate further deviations from normal findings.

Final Selection

The last step in this project was to select the best clinical indicators (now serving as predictor variables) of severe respiratory distress in older OOH patients. This was done using a logistic regression model based on data collected from a retrospective chart review for both predictor and outcome variables. The predictor variables – as determined by the results of the Delphi survey of the OOH experts – were obtained from OOH records, and the binary outcome variable - severe distress vs. not severe distress – was determined through a review of the emergency department (ED) record for that patient visit. The operational definition of severe distress was either a) the physician's opinion that the patient arrived in severe distress, or b) evidence that the patient was placed on bi-phasic or continuous positive airway pressure (BiPAP or CPAP) ventilatory assistance, or was intubated, within an hour of arrival to the ED, or 3) the patient was admitted to an intensive care unit because of the severity of the respiratory problem.

This step required matching the OOH record with the ED record, and OHSU's IRB approved two methods of doing this, depending on the agencies involved. In one instance, the OOH data were obtained directly from the provider agency. In this case, the

agency provided access to records of patients 50 years of age and older who were transported with a respiratory problem as either the primary reason for the call, or for whom the medics identified a respiratory problem. An additional criterion was that these patients were transported to a designated receiving hospital at which the investigator had separate IRB approval to obtain ED chart data. The investigator then obtained access to the records of the ED visit associated with each patient incident, and data regarding outcome variables were abstracted. Samples of the data abstraction forms for both predictor and outcome variables are contained in Appendices 8 and 9, respectively.

With other agencies, direct access to EMS records was not possible. In this case, the investigator worked with ED personnel at one identified hospital to identify patients who met the study criteria, using a query of their ED database. Results of this query were sent to the medical records department, where the full chart for that visit was obtained. The investigator was able to obtain the EMS information from the copies of EMS records that were contained therein; ED information was obtained at the same time.

In neither of these cases was any protected health information (PHI) abstracted from the records for use in the study. All hard copies of lists containing PHI used to match cases were destroyed as soon as the match between the EMS and ED records was complete and data were verified. Electronic files were deleted from the investigator's password protected files as soon as the data collection for those cases was completed. These procedures were all approved by the IRBs at the investigator's institution and at participating hospitals and EMS agencies not affiliated at that institution.

A power calculation was done to estimate the sample size needed for the logistic regression (LR). The sample size was determined by simulations based on the following parameters: an estimated 15% prevalence of severe distress in the population of interest; estimated prevalences of severe scores in five sample predictors of between 20-30%; power of .80; odds ratios (ORs) of 2.0 for significance; between three and five significant predictors out of ten total predictors in the model. Varying these conditions, the simulations resulted in median sample sizes of 325 if five predictors were significant, and 438 if three were significant. The initial sample size for this portion of the study was set at 450 cases.

Human Subjects Considerations

Potential risks to participants

Human subjects were involved in this study in two ways. In the first, the investigator spent time observing and interviewing OOH providers while they worked. Issues of concern in this instance were direct risks to the providers and confidentiality. The risks to the providers were minimal, consisting of psychological discomfort/distress at being observed or answering questions. The investigator worked with the providers to minimize this risk by assuring potential participants that their participation was entirely voluntary, and that even after the investigator arrived to begin a session, she would readily leave if anyone was uncomfortable about being included. Confidentiality was assured by simply not collecting any potentially identifying information from the providers. No names – or any other identifying information - were included in the data

collection. The investigator identified OOH providers by a study number that was not linked in any way to the actual provider.

Human subjects consideration during the chart data abstraction phase was limited to ensuring the confidentiality of the data. This was done by destroying any protected health information that was used to match the OOH medical record with the ED medical record. No protected health information was abstracted on either of the data collection tools.

Participation of women, children, and minorities

Because this study was focused on OOH care of older patients, children (and all others under 50 years of age) were excluded. No special effort was made to recruit women or minorities into this study, just as there was no attempt to exclude them. It was expected that women and minorities would appear in proportion to their percentage of the population in the geographic areas in which the investigator collected her data.

Data Analysis

Item Selection

Literature review

Items identified through the literature review were compiled in a straightforward manner. Many items, e.g. respiratory rate and skin color, were mentioned in all or most sources. Other items, e.g., *pulsus paradoxis* or pursed-lip breathing, were mentioned less

frequently. However, all items mentioned anywhere in the literature review were retained in the initial list.

Review of out-of-hospital records

Items identified through review of OOH records were compared to the initial list obtained from the literature review. Any new assessment parameters that hadn't been identified through the literature review were added to the list of possible predictors.

Observations and interviews with out-of-hospital providers

The investigator reviewed her notes from the observations, debriefings, and interviews to identify the assessment parameters the OOH providers used, or said they used, in evaluating patients in respiratory distress. These parameters were compared to the list being compiled from other sources; items that had not yet been identified were added to the master list.

Item Reduction

Analysis of the results of the Delphi surveys sent to the experts was based on median rankings. Although it is traditional to use means and standard deviations in establishing ranking systems, Bassett and Persky (1999) suggest the median as a more robust estimator of rank. Using the median provides two important properties: a) the rankings reflect the majority of individuals' opinions, not just their scores, and b) a change in ranking of any voter is reflected in a change in overall ranking in the same direction. Median rankings, unlike means, are relatively insensitive to outliers. And, finally, median rankings are a good reflection of the majority of participants.

To determine the median rankings of the first survey, all raters' scores for each variable were summed, arriving at a total score for each variable within each dimension. A table was then constructed in which the variables were ranked first by the total severity score, then by the total reliability score, and finally by the total usability score. Two summary scores were also calculated: the sum of the severity plus reliability scores; and the sum of all three scores. Medians, means, and standard deviations were computed for each of the dimensions; the means and standard deviations being provided for comparison purposes only.

The second survey of experts was analyzed in a similar fashion. In this case, there was only one dimension involved, and scores – total, median, and mean - for each item were summed across raters. The ten predictors with the best (in this case lowest) median scores were included in the final model.

Final Selection

Analysis of data to determine the final list of predictors for the IRD was done using SPSS version 10 (SPSS for Windows, 2000) and R (2003). This analysis includes basic descriptive statistics, including age, gender, and racial distributions; EMS agency; and reasons for activating the EMS system. The frequencies of all predictor and outcome variables were also obtained. The relationships between the predictors, and between the predictor and outcome variables were evaluated using *Chi-square* tests for independence, odds ratios, and correlations using Kendall's *tau*.

A binary logistic regression was then run, with the outcome of severe distress vs. no severe distress, using predictors identified in initial analysis as having some relationship to the outcome variable. Nagelkerke's R^2 was used to evaluate the strength of association of the model (Tabachnick & Fidell, 2001); goodness-of-fit was tested using a deviance lack of fit test based on counted proportions by level of predictor (M. Lasarev, personal communications, July 24, 2003) and the Hosmer-Lemeshow lack of fit test (Collett, 1991; Tabachnick & Fiddell, 2001). These tests actually look at the degree to which there is a *lack of fit* between what the model predicts and what is actually observed. The null hypothesis for these tests is that there is no lack of fit between the model and the observed data. *Non-significant p values* arising from these tests result in retaining the null hypothesis indicating no lack of fit (interpreted as a good fit). A significant *p* value in either lack of fit test would require rejecting the null hypothesis of no lack of fit, thus indicating a poor fit between the model's predictions and the actual data.

The form and fit of the model(s) that resulted were also evaluated using half normal plots of deviance residuals with simulated envelopes (Collett, 1991) as well as doing side-by-side comparisons of the degree to which their predictions of severe distress compared with the actual data.

CHAPTER 4

RESULTS

The chapter contains the results of all phases of this project. The first section contains results of the initial work to select appropriate items for the Index of Respiratory Distress and reduce the list of possible predictors to a number small enough to allow for statistical modeling. Included are results of the initial item selection procedures, reduction of items through the Delphi survey of EMS experts, and confirmation of items through preliminary review of EMS documentation. The first section ends with the operational definitions of the final predictors identified for the analysis.

The second section presents the results of the retrospective review of OOH and emergency department (ED) medical records from OOH patients 50 years of age and older who were treated for respiratory distress. Results of the binary logistic regression investigating the relationship between the identified predictors and the presence of severe respiratory distress are presented in the last portion of this section, along with the discussion of the form and fit of the models that emerged.

Item Selection

Literature Review

The review of the literature provided an initial list of 17 respiratory parameters that were either recommended in standard texts or used in one of the clinical decision rules contained in Appendix 1. These are summarized in Table 5, below.

Table 5

Initial List of Parameter from Review of Literature

Nasal flaring
Tracheal tugging
Retractions (intercostal, supraclavicular)
Accessory muscle use
Paradoxical respiratory movement
Position of patient in bed or chair
Abnormal respiratory rate
Abnormal heart rate
Abnormal skin signs
Altered level or consciousness/mentation
Inability to speak full sentences
Pursed-lip breathing
Audible respiratory sounds
Dyspnea
Pulsus paradoxus
Peak expiratory flow rate (PEFR)
Oxygen saturation (O2 sat)

Observations and Interviews

A total of 108 hours was spent riding with OOH providers from three agencies in the Portland metropolitan area. The investigator rode during a variety of hours, covering the entire 24-hour day except for the hours between 3:00 am and 7:00 am. Rides took place in the inner urban core, the suburbs, and some hours in semi-rural areas. During this observation time, 18 patient contacts were observed, only five of which involved assessing the respiratory status of a patient; none of these patients was in anything more than mild respiratory distress. Providers who cared for these patients were debriefed to ascertain what they had assessed to reach their conclusions regarding those patients' levels of distress. In addition, 13 providers with from less than one year to greater than 25

years of experience were interviewed in depth regarding their assessments of older patients in respiratory distress. A summary description of OOH providers interviewed in depth, and the relevant ALS patient contacts appears in Appendices 10 and 11, respectively.

Table 6 contains a summary of the data collected during the rides with these OOH providers. The table lists the clinical indicators identified by the providers, and the number of providers in each agency who identified each indicator as important in assessing older patients in respiratory distress.

Table 6

Summary of Observation and Interview Data from Ride-Alongs with OOH Providers

Parameter	Agency 1[a]	Agency 2	Agency 3	Total
Pre-arrival information[b]	1	2	0	3
Position of patient	2	3	4	9
Audible breath sounds	1	2	1	4
Fatigue, "looking pooped"	1			1
"Looking pale," or other abnormal skin color	2	3	3	8
Unable to speak	2	1	2	5
Respiratory rate and quality[c]	1	3	2	6
SaO2 (oxygen saturation)	2	2	2	6
Diaphoresis	1	1		2
Respiratory effort	2		2	4
Nasal flaring	2			2
Looking anxious	1	2	1	4
Decreased level or consciousness	1	1		2
Decreased exercise tolerance	1	1		2
Dyspnea[d]- new or worse compared to usual for pt.	1	1	1	3
Agitation			2	2
Intercostal/accessory muscle use		1	1	2
Skin temperature	1	1		2

[a] Number in the column indicates number of times OOH providers in these agencies identified this as an important indication of severity of respiratory distress.

[b] Although not a clinical indicator, this parameter was mentioned enough to be included as a possible predictor variable.

[c] Providers specified they did not count respirations *per se*, but used an overall impression of "too fast," or "rapid and shallow."

[d] The providers used this word on their own without prompting from the investigator.

Final List of Possible Predictors

Table 7 is a list of the indicators identified in both sources, i.e., the literature review and the rides with OOH providers, identified by the source of each indicator.

Table 7

Sources of Possible Predictor Variables

Predictor	**Source**[a]
Nasal flaring	B
Tracheal tugging	L
Retractions (intercostal, supraclavicular)	L
Accessory muscle use	B
Paradoxical respiratory movement	L
Position of patient in bed or chair	B
Abnormal respiratory rate	B
Abnormal heart rate	L
Abnormal skin signs	B
Altered level or consciousness (LOC)/mentation	B
Inability to speak full sentences	B
Pursed-lip breathing	L
Audible respiratory sounds	B
Dyspnea	B
Pulsus paradoxus	L
Peak expiratory flow rate (PEFR)	L
Oxygen saturation (O2 sat)	B
Pre-arrival information	R
Fatigue, "looking pooped"	R
Respiratory effort	R
Decreased exercise tolerance	R
Agitation	R

[a] L=literature review; R = ride-along (includes observation, debriefing, and interviewing OOH providers); B=both

As can be seen, there were many items identified in both the literature review and during the observations and interviews; others were found in one source but not the other. All were included in the list of possible predictors for the item reduction stage of this

study, except "pre-arrival information;" that was removed as it is not a direct observation of patient status. In the list of clinical indicators used in the Delphi survey, the components of vital signs, i.e., blood pressure, heart rate, and respiratory rate, were listed separately; this was also done for the components of skin signs, i.e., skin temperature, color, and moisture. Table 8 shows the list of clinical indicators that was used in the Delphi survey sent to EMS experts for their input into the item reduction process.

Table 8

List of Clinical Indicators of Severity of Respiratory Distress in Older Adults Used in Delphi Survey of EMS Experts

Possible Predictors
Nasal flaring
Tracheal tugging
Retractions
Accessory muscle use
Paradoxical respiratory movement
Position of patient when provider arrives - (Relaxed/supine vs. semi-fowlers vs. high fowlers/tripod)
Altered blood pressure
Altered respiratory rate
Altered heart rate
Abnormal skin temperature
Abnormal skin color
Abnormal skin moisture
Level or consciousness/mentation
Inability to speak full sentences
Pursed-lip breathing
Audible respiratory sounds
Dyspnea
New or significantly increased from baseline
Fatigue
Decreased oxygen saturation
Compromised peak expiratory flow rate
Increased respiratory effort
Facial expression indicating anxiety/stress/panic
Agitation
Decreased exercise tolerance
Pulsus paradoxus

Item Reduction

Results from Delphi Survey of EMS Experts

Six of the eight EMS experts identified by the investigator completed and returned the initial survey. These six included two physicians, two nurses, and two paramedics. One paramedic and one physician did not return their first survey, and they were dropped from further participation. The current positions and previous experiences in EMS work for the six remaining experts are summarized in Appendix 12 .

The experts rated these clinical indicators on three dimensions: severity, or the importance of the item as an indicator of disease severity; reliability, or the ability of OOH providers to assess the indicator consistently; and usability, or the extent to which providers already use this assessment in their current practice. Each indicator was to be rated on a scale of 1-3, with 1 indicating the item was not very important, reliable, or useable, and 3 indicating it was highly important, reliable, or useable. The results of the experts' rankings were tabulated and total scores were derived for each item in each dimension. For example, *nasal flaring* received a total score of 12 for severity, 12 for reliability, and 11 for usability. The full results of this tabulation appear in Appendix 13.

The scores were then added sequentially, first severity and reliability were summed; usability was then added. The range of possible scores was 6-18 for ratings in a single dimension; 12-38 for two dimensions; 18- 54 for all three dimensions. As described in detail in Chapter 3, median rather than mean scores are used in these analyses as their numerical properties are better suited to rankings. Medians were

calculated for each dimension and for both sums; means are provided for comparison purposes. These are shown in Table 9.

Table 9

Summary of First Delphi Survey Rankings on Three Dimensions

Predictor Variable	Total Severity Score	Total Reliability Score	Total Usability Score	Total Severity + Reliability	Total All Dimensions
LOC/mentation	17	17	17	34	51
Respiratory rate	15	18	18	33	51
Inability to speak full sentences	15	18	15	33	48
Pursed lip breathing	15	14	11	29	40
Retractions	15	13	13	28	41
Decrease SaO2	14	18	18	32	50
Position upon arrival	14	14	16	28	44
Accessory muscle use	14	14	13	28	41
Skin temp	14	14	13	28	41
Dyspnea	14	12	15	26	41
Compromised PEFR	14	9	7	23	30
Increased resp. effort	13	13	15	26	41
Fatigue	13	10	10	23	33
Paradoxical resp. movement	13	8	8	21	29
Audible resp. sounds	12	16	16	28	44
Skin moisture	12	14	16	26	42
Expression of anxiety/stress/ Panic	12	12	12	24	36
Nasal flaring	12	12	11	24	35
Pulsus paradoxus	12	8	6	20	26
Pulse rate	11	18	18	29	47
Agitation	11	13	14	24	38
Tracheal tugging	11	9	8	20	28
Blood pressure	9	17	18	26	44
Skin color	9	13	17	22	39
Decreased exercise tolerance	9	6	6	15	21
Median scores	**13**	**13**	**14**	**26**	**41**
Mean	**12.8**	**13.2**	**13.24**	**26**	**39.24**

The process of choosing clinical indicators for the second and final survey was done based on the median scores. Using sums of dimensions rather than any single dimension accounted for circumstances in which an item may be a good indicator of physiological compromise but may also be difficult for OOH providers to obtain consistently or reliably. Further, the item may also be something OOH providers do not use routinely in their practice. For example, peak expiratory flow rate (PEFR) had a total score of 14 (out of a possible 18) in severity, but only 9 in reliability, and 7 in usability. The sum of severity and reliability was 23, 3 points (11.5%) below the median score. When the usability score was added, the total, 30, was 11 points (26.8%) below the median. Considering all three dimensions, PEFR was not chosen for the second survey.

Using this process for all items, the original list of 25 predictors was reduced to the following 15 clinical indicators: retractions, accessory muscle use, position of patient when provider arrives, altered blood pressure, altered respiratory rate, altered heart rate, abnormal skin temperature, abnormal skin moisture, level of consciousness/mentation, inability to speak full sentences, pursed-lip breathing, audible respiratory sounds, dyspnea, decreased oxygen saturation, and increased respiratory effort.

All six experts who completed the first survey also completed the second. Although they were asked specifically not to rank any two items the same level, and were given clear opportunities to either add or remove items apart from the ranking task, three experts either omitted ranking certain items, inserted and ranked an item not on the list, or removed an item from the list, ranking only the 14 remaining. In order to preserve as

much information as possible, the investigator imputed ranks for these three experts in the following ways:

1. In the first instance, one expert removed "abnormal skin temperature" indicating he didn't think it was an important indicator of physiological compromise. At the same time, he inserted "peak flow" and gave it a ranking of 4. To remedy this situation, the investigator assigned a value of 15 to "abnormal skin temperature," removed "peak flow" and adjusted the rankings accordingly (what had been ranked 5 became 4, 6 became 5, etc.).
2. In the second instance, the expert removed "increased respiratory effort" from his list, indicating it was "a composite of other items," leaving him with only 14 items to rank. In this case, the investigator assigned a rank of 15 to the item.
3. In the last case, the expert omitted ranking two items: "increased respiratory effort" and "altered respiratory rate", indicating he felt these to be the same as "dyspnea", which the expert ranked as 7. To remedy this situation, the investigator took the two items that were unranked, and the item marked as 7, and gave all three items a ranking of 8, the average of 7, 8, and 9, which is how those items should have been ranked if the respondent had followed directions.

After making these adjustments in rankings, the rankings for each indicator were summed across raters. Total summed ranks ranged from 22-73 (possible range 6-90),

with a median of 48, mean of 47.6, and standard deviation of 19.23. In addition, the median score, mean, and standard deviation were calculated for each clinical indicator. Once again, median scores, rather than means, were used for this analysis; means and standard deviation are provided for comparison purposes. The list of indicators was then sorted by median score, the lowest scores indicating the best predictors. Appendix 14 contains the complete spreadsheet for this analysis; Table 10 shows the total scores, median rankings, mean rankings and standard deviations.

Table 10

Summary Scores of Second Rankings From Delphi Survey

Predictor	Total Rank Score	Median Rank	Mean Rank	Standard Deviation
LOC/mentation	22	2	3.67	3.39
Inability to speak full sentences	23	2.5	3.83	3.82
Position of patient	24	3.5	4	2.61
Decreased SaO2	30	4	5	2.68
Accessory muscle use	34	6	5.67	2.501
Dyspnea	45	7	7.5	2.74
Increased resp. effort	48	7	8	3.58
Pursed-lip breathing	49	8	8.17	3.06
Altered Resp. rate	44	9	7.33	3.98
Retractions	49	9	8.17	3.26
Audible respiratory sounds	52	10	8.67	3.89
Abnormal skin moisture	68	12	11.33	1.51
Altered BP	70	13.5	11.67	4.97
Altered heart rate	73	13.5	12.17	2.91
Abnormal skin temperature	83	14	13.83	0.75

The ten best predictors, that is, the predictors with the ten lowest median scores are those with median rankings below 10. Based on these results, the ten clinical indicators considered for use as predictors in the logistic regression model were level of consciousness/mentation; inability to speak full sentences; position of the patient;

decreased oxygen saturation; accessory muscle use; dyspnea; increased respiratory effort; pursed-lip breathing; altered respiratory rate; and retractions.

Confirmation of predictors in OOH medical records

One critical consideration to including any clinical indicator as a predictor in the final model is that it be one that OOH providers are documenting in the patients' records. Table 11 shows the results of the review of 28 de-identified records from one participating EMS agency as they relate to each of the predictors being considered for inclusion in the model.

Table 11

Predictors Found in Review of Records from One Participating EMS Agency[a]

Predictor	Observations from EMS Records
LOC/mentation	On electronic form and noted in narrative, e.g. "A & O x 4" [b]
Inability to speak full sentences	Noted frequently in narrative, e.g., "4-6 word sentences"
Position of patient	Noted frequently in narrative, e.g. "arrived to find patient sitting on chair..."
Decreased oxygen saturation	Mandatory field; providers enter value and level of oxygen delivery, e.g. "94% NRB @10 L/min"[c]
Accessory muscle use	Noted occasionally in narrative, e.g. "accessory muscle use"
Altered respiratory rate	Mandatory field; providers enter value
Dyspnea	Noted once in narrative, e.g. in check box for "Protocol/Assessment" one provider entered, "dyspnea/poss. Pneumonia"
Increased respiratory effort	On electronic form. Choices include normal; increased, not labored; increased and labored or decreased and fatigues; absent; not assessed.
Retractions	Not noted
Pursed-lip breathing	Not noted

[a] Taken from same agency as example in Appendix I

[b] Alert and oriented in 4 dimensions: time, place, person, and activity

[c] Oxygen saturation is 94% with patient on a non-rebreather mask at 10 liters of oxygen per minute flow rate

This process indicated that certain possible predictors such as pursed-lip breathing, retractions, and dyspnea would appear so seldom on the OOH record as to make it impossible to include them in the logistic regression model. In addition, dyspnea appears to be used as a form of field diagnosis rather than a descriptor of patient symptoms. Eliminating these predictors left the following possible predictors for the regression model: LOC/ mentation, inability to speak full sentences, position of patient, decreased oxygen saturation, accessory muscle use, altered respiratory rate, and increased respiratory effort.

Item Scaling and Scoring

Predictor Variables

All items except "accessory muscle use" were scaled 0, 1, or 2, where 2 represents normal findings or no change from the patient's baseline; 1 indicates some change from normal or baseline, but not markedly so; and 0 means radical departure from normal or baseline. Accessory muscle use was scored as either 2 (none noted) or 0 (noted). The operational definition for each score is provided in Table 12.

It should be noted that these definitions were synthesized from the literature on respiratory distress as filtered through the investigator's twenty seven years as a practicing emergency department nurse. The limits for each level represent the investigators best efforts to standardize the interpretation of the range of results commonly recorded when OOH providers assess these various clinical indicators.

Table 12

Operational Definitions for Each Level of Predictor Variable

Predictor	Scale	Description
Level of consciousness/mentation	2	Fully awake, oriented, and cooperative (no change from baseline).
	1	Awake, mildly agitated or disoriented, cooperates only with constant reminding (slightly worse than baseline).
	0	Very agitated, disoriented and uncooperative, **or** unconscious/unresponsive (markedly worse than baseline).
Position of patient	2	Found lying flat or sitting or reclining comfortably (semi-Fowler's) (if bed-bound, no need to alter position).
	1	Found sitting up or high Fowler's (if bed-bound, position altered to aid breathing).
	0	Tripod position or patient sitting fully upright and leaning forward ("sniffing position")
Ability to speak full sentences	2	Speaking in full sentences or more than six-word sentences
	1	Some restriction in ability to speak full sentences; can speak up to six words in one breath
	0	Cannot speak or speaking up to two-word sentences
Decreased oxygen saturation	2	Oxygen saturation 95% or better on room air
	1	Oxygen saturation 95% or better with supplemental oxygen
	0	Cannot maintain oxygen saturation of 95% or better despite delivery of oxygen
Accessory muscle use	2	Not using accessory muscles
	0	Accessory muscle use noted
Altered respiratory rate	2	Respiratory rate between 8 and 20 per minute
	1	Respiratory rate between 6-8 or 20-32 per minute
	0	Respiratory rate less than 6 or more than 32 per minute
Increased respiratory effort	2	Normal
	1	Rate increased but not labored
	0	Rate increased and labored, or patient apneic

The Outcome Variable

The outcome variable is *severe respiratory distress*. Severe respiratory distress is defined operationally in one of three ways:

1. There is a statement by an ED physician that the patient arrived in severe distress, or an equivalent phrase.
2. The patient was placed on assisted ventilation (continuous or bi-phasic positive airway pressure or endotracheal intubation) within one hour of arrival to the ED.
3. The patient was admitted to the ICU for treatment and monitoring of the respiratory problem.

The presence of any one of these findings in the record was considered sufficient for a positive finding of severe distress.

Results of Logistic Regression

Data for the predictor variables were obtained from a retrospective chart review of OOH medical records from three EMS agencies. Data regarding the outcome variable, *severe distress* were obtained from the ED medical record for that episode of care. Again, Appendices 8 and 9 contain the tools used to abstract the information from the OOH and ED medical records respectively.

Sample

A convenience sample OOH and matching ED medical records was obtained for patients 50 years of age and older who received OOH care for respiratory problems between January 1, 2001 and June 30, 2002. The sample was obtained using two different procedures. In the first instance, one EMS agency identified records of appropriate patients, and data on the predictors were abstracted from those records. These were subsequently matched to the ED visit for that episode, and data regarding the outcome variable were then abstracted. In the second instance, complete hospital records for appropriate patients were obtained, and data on both the predictors and outcome were abstracted from OOH and ED records that were retained as part of the patient's permanent record. These two procedures yielded different numbers of charts, and exclusion criteria were slightly different for each procedure.

Exclusion criteria for the first sampling procedure are summarized in Table 13. Using this procedure, the EMS agency identified 504 records. After applying the exclusion criteria, 280 remained in the database.

Table 13

Reasons for Excluding EMS Cases: Out-of-Hospital Accessed Data

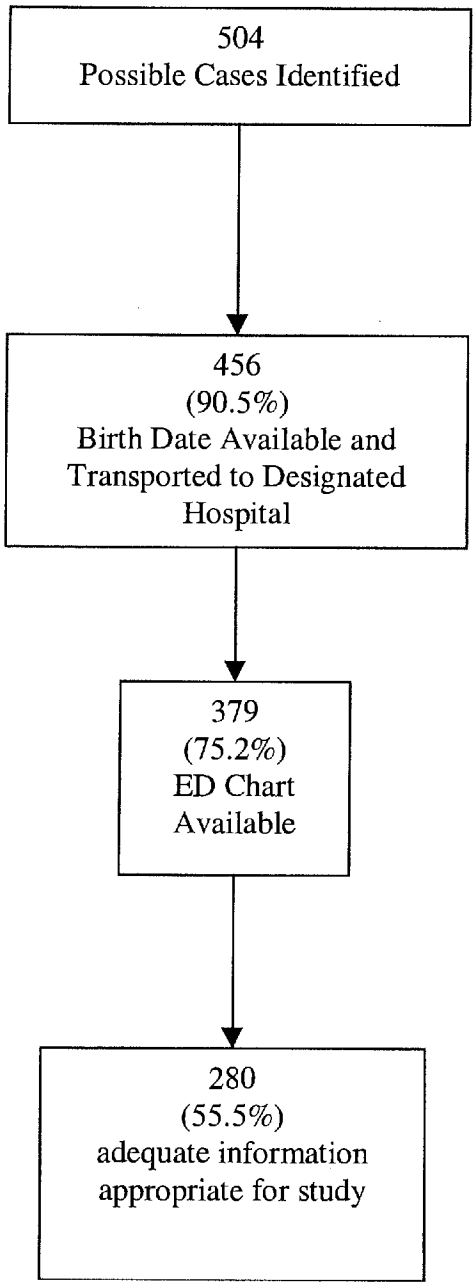
Reason for Excluding Cases	**Number of Cases Excluded** [a]
Fewer than 4 parameters noted in EMS record	77
No ED record match found	77
No date of birth on EMS record – no ED match possible	25
Patient transported to other than target hospital	20
Case inappropriate for study	14
Field intubation	6
Other reason, e.g., patient was direct admit	8

[a] Several cases appear in more than 1 category.

Figure 3 presents a graphic display of the final sample from this source.

Figure 3: Sample of Matching EMS and ED Records

Source: EMS Agency – EMS data from EMS agency; ED data from hospital record



The second procedure required identifying hospital medical records of eligible patients. The medical information system did not allow for direct query of their database; at their suggestion, the hospital financial records were used initially. This query was based entirely on discharge diagnosis and did not allow for selection based on the patient's mode of arrival to the ED. As a result, only 13% (n=13) of the initial cases identified fit the criteria for the study, and only 4 of those had useable data. A second approach relied on the ED database to select patients based on their complaint upon arrival, with ability to screen for age and mode of arrival. This procedure proved to be more satisfactory. Together these two queries resulted in identifying 136 cases between January 1 and December 31, 2001. Table 14 describes the reasons for excluding charts from this sample; 31 cases remained in the sample from this source. Figure 4 illustrates how this final sample was derived.

Table 14

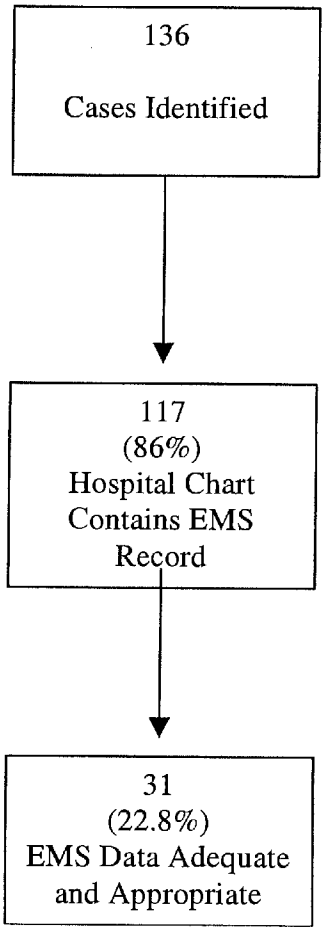
Reasons for Excluding Cases: Hospital Accessed Data

Reason for Excluding Cases	**Number of Cases Excluded**
EMS records lack adequate data[5]	79
No EMS record on chart	19
Case inappropriate for study, e.g., patient seen for non-respiratory problem	7

[5] The copies of EMS records in the hospital charts were sometimes incomplete; one agency left only the narrative page of their official two-page record. The missing page was formatted for computer scanning, but contained information such as patients' vital signs, level of respiratory effort, Glasgow Coma Scale, etc. that may or may not have also been included on the narrative page.

Figure 4: Sample of Matching EMS and ED Records

Source: Hospital Chart: EMS and ED Records both from Hospital Chart



The total sample from both sources was 311 cases of matched EMS and ED records with data on at least four of the seven predictor variables. Of those, 124 cases had complete data on all seven predictor variables.

Characteristics of the sample

Table 15 summarizes the characteristics of the analytic sample and compares it to the full sample. None of the differences noted between the analytic sample and full sample was statistically significant except age; the analytic sample was slightly older ($t_{(292)}$ =-2.18, p=.030; adjusted for unequal variances). The vast majority of cases came from a single EMS agency; these cases were obtained directly from the EMS records, with follow-up at the designated receiving hospital, as was shown in Figure 3. Virtually all the calls were for a primary respiratory complaint. The sample was primarily European American, but race/ethnicity was indicated as "unknown" in a substantial minority of cases. There was no significant difference in mean ages between the men and women in the analytic sample ($t_{(103)}$ =-1.75, p=.075; adjusted for unequal variances).

Table 15

Characteristics of Sample

Characteristic	Analytic Sample n =124 (100%)	Total Charts Reviewed n=311 (100%)
Gender		
Men	55 (44.4)	122 (39.2)
Women	69 (55.6)	189 (60.8)
Age		
Mean	76.2	74.5
Standard deviation	9.92	11.19
Median	77	76
Race/ethnicity		
European-American	94(75.8)	224 (72.0)
African-American	3(2.4)	6 (1.9)
Asian	1(0.8)	4 (1.3)
Other	5(4.0)	11 (3.5)
Unknown	21(16.9)	66 (21.2)
Reason for 9-1-1 call		
Primary respiratory	117(94.4)	295 (94.9)
Secondary respiratory	3(2.4)	4 (1.3)
Not respiratory	3(2.4)[a]	~~66 (21.2)~~[b] 10(3.2)
EMS agency		
1	11(8.9)	45 (14.5)
2	2(1.6)	3 (0.9)
3	110(88.7)	260 (83.6)
4	1(0.8)	3 (0.9)

[a] Reason for call was missing on one case

[b] Reason for call was missing on two cases

Frequency of predictor and outcome variables

Table 16 shows the frequency of the seven predictor variables and the binary outcome variable among the 124 cases. All levels of all seven predictors were present in

the sample. Level of consciousness had the least variability, with 89.5% of the sample rated as having a normal level of consciousness, and respiratory effort was the most evenly distributed. The frequency of severe distress in this sample was 20.2% (25 cases).

Table 16

Frequency of Predictor and Outcome Variables

Variable	N=124 (100%)
Predictors	
Level of consciousness	
Normal	111(89.5)
Mildly alterated	9(7.3)
Extremely altered	4(3.2)
Ability to speak	
>6 words-full sentences	75(60.5)
3-6 word sentences	30(24.2)
0-2 word sentences	19(15.3)
Accessory muscle use	
None noted	90(72.6)
Accessory muscle use noted	34(27.4)
Position of patient on arrival of EMS providers	
Flat or unremarkable	41(33.1)
Head elevated, sitting comfortably	57(46.0)
Bolt upright; tripod; "sniffing position"	26(21.0)
Oxygen saturation	
Equal to or greater than 95% on room air	25 (20.2)
Equal to or greater than 95% on supplemental oxygen	79(63.7)
Less than 95% on 100% oxygen	20 (16%)
Respiratory rate	
Between 8 and 20 breaths per minute	23(18.5)
Between 6-8 or 20-32 breaths per minute	67(54.0)
Fewer than 6 or more than 32 breaths per minute	34(27.4)
Respiratory effort	
Normal rate and depth	43(34.7)
Increased rate, not labored	41(33.1)
Increased rate and labored	40(32.3)
Outcome[1]	
Not in severe respiratory distress	99(79.8)
In severe respiratory distress	25(20.2)

[1] Patients were determined to be in severe respiratory distress if any of the following was found in the ED chart: a) a statement indicating the patient arrived in severe respiratory distress, or b) the patient was intubated or placed on assisted ventilation within one hour of arrival in the ED, or c) the patient was admitted to an intensive care unit for treatment of the respiratory problem.

Preliminary Analyses

Table 17 shows the correlation between the seven predictor variables and between the predictors and the outcome variable, severe distress. All predictors were correlated with the outcome variable in the expected (negative) direction, and with the exception of the correlation between *level of consciousness* and *position of patient*, all were positively correlated with each other. Table 18 summarizes the results of the univariate analyses using both Chi-square and odds ratios. Chi-square tests indicated there may be some relationship between five of the predictors and severe distress (level of consciousness and oxygen saturation failed to achieve statistical significance). Odds ratios using the "normal" category as referent are greater than 2.0 for all categories except mild alteration in level of consciousness. The analyses suggest the possibility of significant relationships between each of the predictors and the outcome; therefore, the decision was made to include all seven predictors in the logistic regression model.

Table 17

Correlation Between Predictors and Outcome Variables (N=124)

	LOC[a]	Speak[b]	Muscle[c]	Position[d]	Saturation[e]	Rate[f]	Effort[g]	Severe[h]
LOC	**1.000**							
Speak	.227	**1.000**						
Muscle	.030	.546	**1.000**					
Position	-.051	.418	.399	**1.000**				
Saturation	.189	.326	.239	.205	**1.000**			
Rate	.186	.475	.439	.269	.301	**1.000**		
Effort	.107	.554	.496	.410	.344	.568	**1.000**	
Severe	-.032	-.486	-.502	-.334	-.193	-.308	-.431	**1.000**

[a] LOC = level of consciousness; [b] Speak = ability to speak full sentences; [c] Muscle = accessory muscle use
[d] Position = patient's position in bed or chair; [e] Saturation = oxygen saturation; [f] Rate = respiratory rate
[g] Effort = respiratory effort; [h] Severe = patient in severe distress (outcome measure)

Table 18

Chi-square and Odds Ratios for Predictors and Severe Distress

Predictor Variable	Chi-square (degrees of freedom)[a]	*p* value	Odds Ratio[b]
Level of consciousness	2.68(2)	.262	
Mildly altered			0.5
Extremely altered			4.0
Ability to speak	31.95(2)[c]	<.001	
3-6 word sentences			16.67
0-2 word sentences			27.78
Accessory muscle use	31.27 (1)	<.001	
Accessory muscle use noted			14.06
Position of patient on arrival of EMS providers	19.26(2)	<.001	
Head elevated, sitting up			2.34
Bolt upright; tripod; "sniffing position"			12.5
Oxygen saturation	5.03(2)[d]	.08	
Equal to or greater than 95% on supplemental oxygen			2.82
Less than 95% on 100% oxygen			6.0
Respiratory rate	12.92(2)[e]	.002	
Between 6-8 or 20-32 breaths per minute			11.26[f]
Fewer than 6 or more than 32 breaths per minute			31.3
Respiratory effort	26.43(2)	<.001	
Increased rate, not labored			22.5[f]
Increased rate and labored			82.2

[a] Chi-square was computed for each predictor variable as a whole, not separately for each level within the predictor.

[b] Odds ratios were computed for each level within the predictor using normal findings (coded "0") as the referent.

[c] One cell had an expected count of 3.8.

[d] One cell had an expected count of 4.

[e] One cell had an expected count of 4.64.

[f] There were no observed cases in which patients with normal findings were also in severe distress; 0.5 was added to each cell to compute odds ratios.

Initial Regression Model

The initial logistic regression was run treating all predictors as categorical, with normal findings as the referent category. The results of this first step are shown in Appendix 15. Using backward extraction, with criteria p=.05 for entry and .10 for removal, three predictors were identified: accessory muscle use, respiratory effort, and ability to speak full sentences. These categories were then entered into a reduced model; these results are summarized in Table 19.

Table 19

Results of Logistic Regression (N=124)

	Estimate	**SE**	**z-value**	***p***
Muscle[a]	1.197	0.6054	1.977	.048
Effort 1[b]	7.986	17.954	0.440	.660
Effort 2	8.546	17.955	0.476	.634
Speak 1[c]	1.407	0.777	1.812	.070
Speak 2	1.654	0.829	1.995	.046

R^2 = .492; Deviance lack of fit (DLOF)= $\chi^2_{(9)}$= 11.183, p=.26; Hosmer –Lemeshow (HL) = $\chi^2_{(8)}$= 3.182, p=.874

[a] Accessory muscle use present

[b] Effort = respiratory effort; 1=mild alteration; 2=severe alteration

[c] Speak= ability to speak full sentences; 1=mild alteration; 2=severe alteration

The R^2 in this model is nearly identical to that in the full model, .512, ($\chi^2_{(2)}$=2.272; p=.972). Both lack of fit tests indicate the model fits the data well.

Exploring Alternative Models

The model without accessory muscle use

The investigator was interested in exploring whether it was necessary to retain two problematic predictors, accessory muscle use and respiratory effort. In the case of accessory muscle use, the investigator noted that over one-third of the original 311 cases were missing data (n=110, 35.4%) this suggests the predictor is not recorded often enough to make it a useful predictor in a clinical decision rule. In addition, accessory muscle use is barely statistically significant. What was found was that although a model with just respiratory effort and ability to speak continued to fit the data fairly well, there was a statistically significant reduction in explanatory power when the R^2 drops from .492 to .457 ($\chi^2_{(1)}$=6.288; p=.045). Table 20 presents the model without accessory muscle use included.

Table 20

Logistic Regression with No Accessory Muscle Use Data N=124

	Estimate	**SE**	**z-value**	***p***
Effort 1[a]	8.077	17.850	0.453	.651
Effort 2	8.977	17.850	0.504	.614
Speak 1[b]	1.877	0.727	2.595	.009
Speak 2	2.063	0.793	2.602	.009

R^2=.457; DLOF = $\chi^2_{(3)}$= 0.918, p=.82; HL= $\chi^2_{(8)}$= 6.184, p=.627

[a] Effort = respiratory effort; 1=mild alteration; 2=severe alteration

[b] Speak= ability to speak full sentences; 1=mild alteration; 2=severe alteration

The model without respiratory effort

The models both with and without accessory muscle use retained as a predictor exhibit extremely large standard errors for respiratory effort; in addition, the *p* values indicate that respiratory effort is not a statistically significant predictor. The investigator was, therefore, interested in determining the importance of respiratory effort in the model. Removing respiratory effort, while it resulted in more reasonable standard errors and greater statistical significance for the remaining predictors, again resulted in a significantly reduced R^2 ($\chi^2_{(2)}$=6.364; *p*=.042). In addition, the DLOF test is significant, indicating this model may fit poorly, although the HL test remains non-significant. The DLOF test is the more trustworthy of these two tests in this case, where the variables are treated as categorical rather than continuous.

Table 21

Logistic Regression with No Respiratory Effort Data (N=124)

	Estimate	**SE**	**z-value**	***p***
Muscle[a]	1.561	0.596	2.622	.009
Speak 1 [b]	2.020	0.767	2.641	.008
Speak 2	2.496	0.815	3.064	.002

R^2=.435; DLOF = $\chi^2_{(2)}$= 7.034, *p*=.03; HL= $\chi^2_{(3)}$= 5.084, *p*=.166

[a] Accessory muscle use present

[b] Speak= ability to speak full sentences; 1=mild alteration; 2=severe alteration

Further Evaluating the Form and Fit of the Models

Half-normal plots of deviance residuals

Half-normal plots of deviance residuals with simulated envelopes provide a good visual indication of how well a reduced model fits the data (Collett, 1991). Appreciating the value of this plot depends on understanding the concepts of deviance residuals, half-normal plots, and simulated envelopes.

The deviance of a particular model is the discrepancy between a full, or saturated model, and the current model under investigation. A saturated model has as many parameters as there are data points and necessarily fits the data perfectly, but is no simpler than the data itself. A reduced model, with a limited number of parameters being fitted, provides a useful summary of the data but at the cost of a perfect fit.The contribution of an individual observation to the total discrepancy between the saturated and reduced models is termed the *deviance residual*, and may take either a positive or negative value. The sum of squared deviance residuals always equals the total deviance between the saturated and reduced models (M. Lasarev, personal communication, July 16, 2003).

The absolute values of the deviance residuals can be plotted against the expected absolute value of a standard normal deviate. The form of this distribution, in which the negative half of the normal distribution is reflected over the positive half is that of half the normal curve; hence its name, a *half-normal plot.*

A *simulated envelope* enhances the half-normal plot of the absolute values of the deviance residuals from the original data set by overlaying the median and 95% confidence band from 19 simulated data sets having the same response probabilities as observed in the sample.Using 19 simulated data sets– in addition to the original data set - results in only a 5% chance that the largest absolute residual from the original data set will lie outside the envelope, and thus the envelope can also be used to assess for outliers (Collett, 1991).

If the model in question is adequate, the plot reveals deviance residuals randomly distributed close to the simulated median line that rises to the right, but all within the 95%confidence band. The data points should not show any systematic deviations from the median line, and occurrence of several points near to or outside the envelope indicates that the fitted model may not be appropriate (Collett, 1991) .

Figures 5-8 show the half-normal plots of deviance residuals with simulated envelopes for the full model containing all seven original predictors, the reduced model with three predictors, the reduced model with accessory muscle use removed, and the reduced model with respiratory effort removed.

Figure 5

Half Normal Plot of Deviance Residuals with Simulated Envelope for Full Model

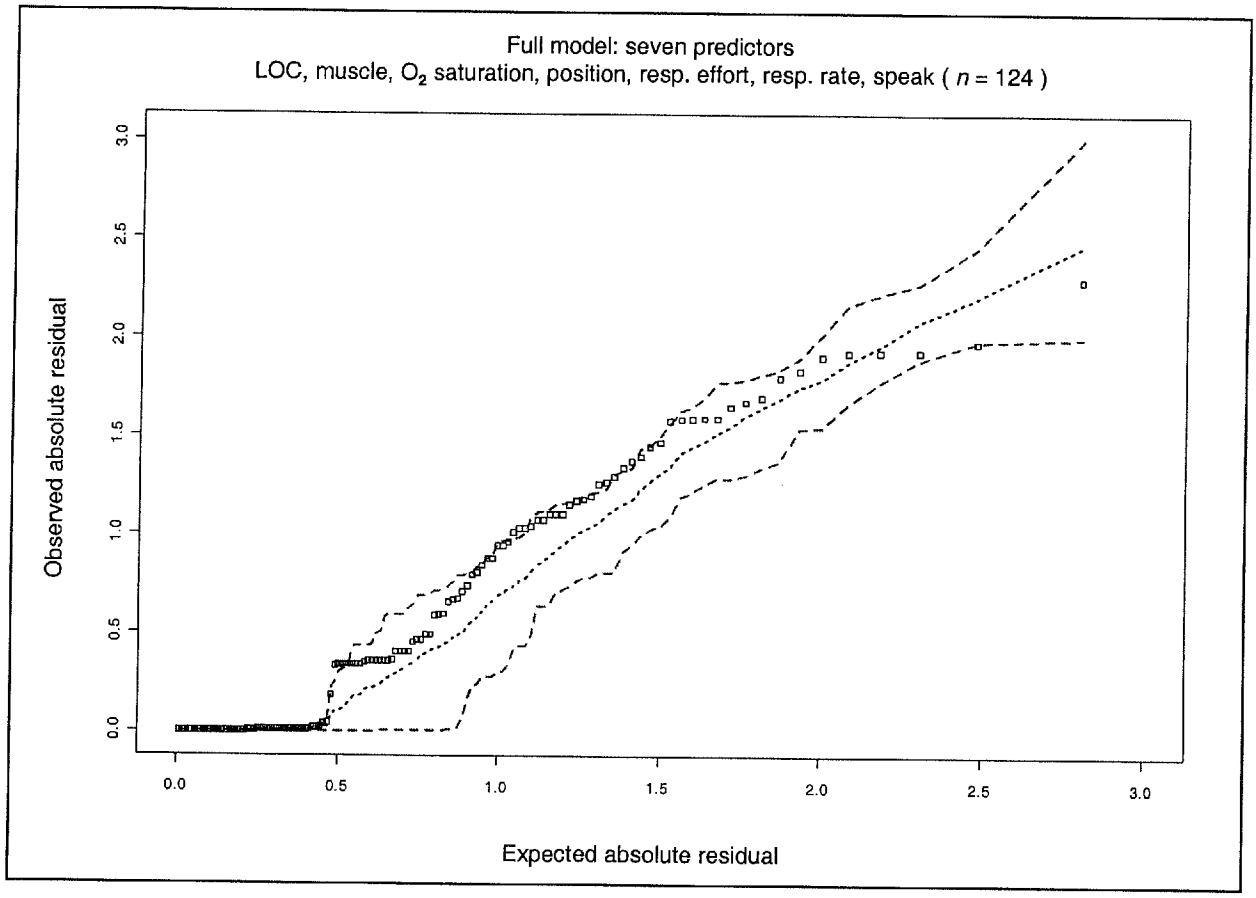


Figure 6

Half Normal Plot of Deviance Residuals with Simulated Envelope for Reduced Model with Three Predictors: Accessory Muscle Use, Ability to Speak, Respiratory Effort

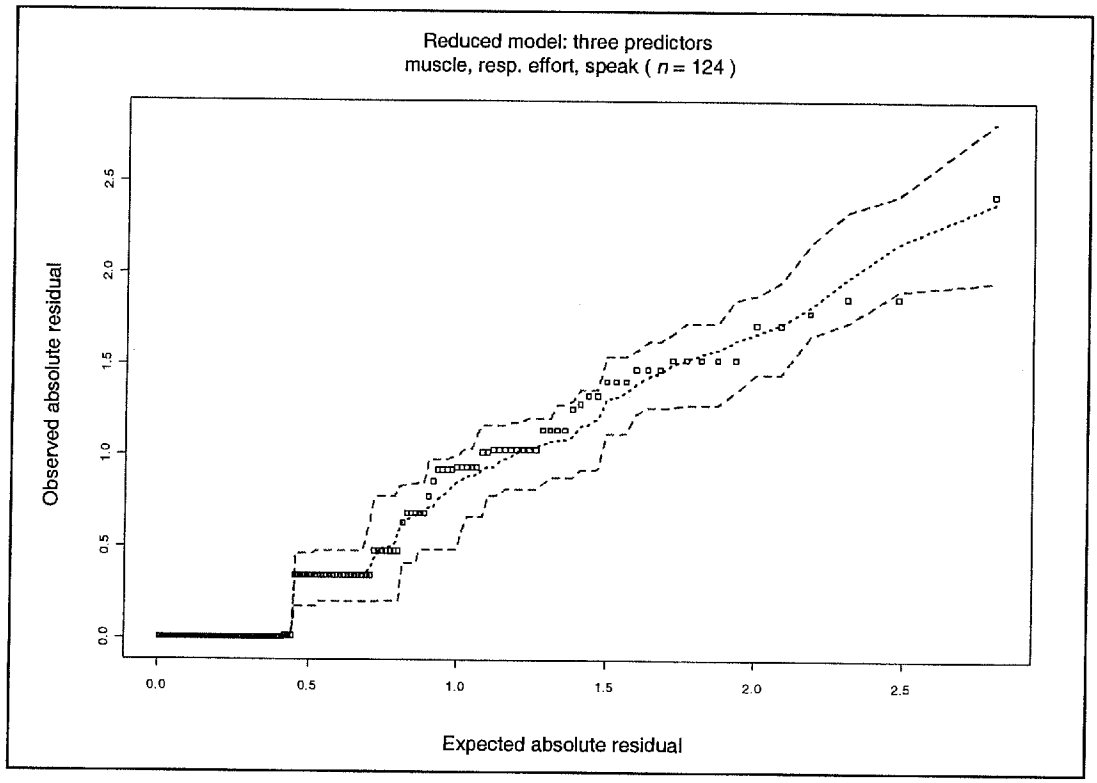


Figure 7

Half Normal Plot of Deviance Residuals with Simulated Envelope for Reduced Model with Two Predictors: Respiratory Effort, Ability to Speak

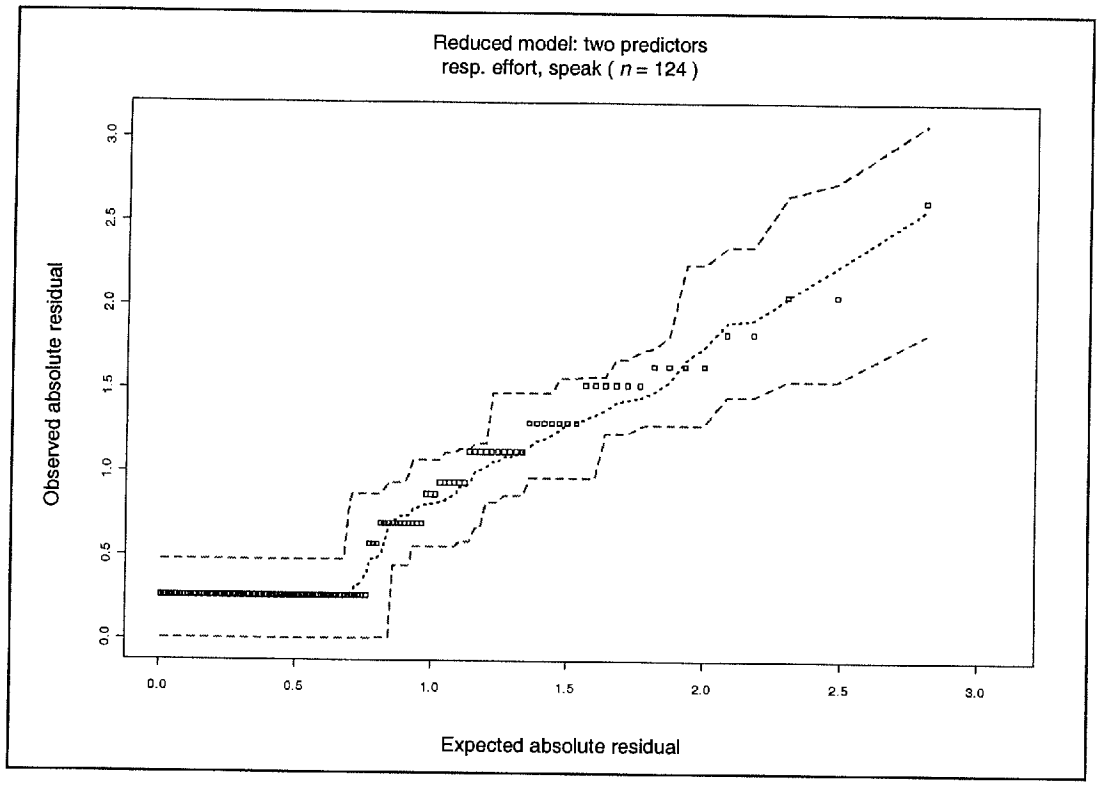
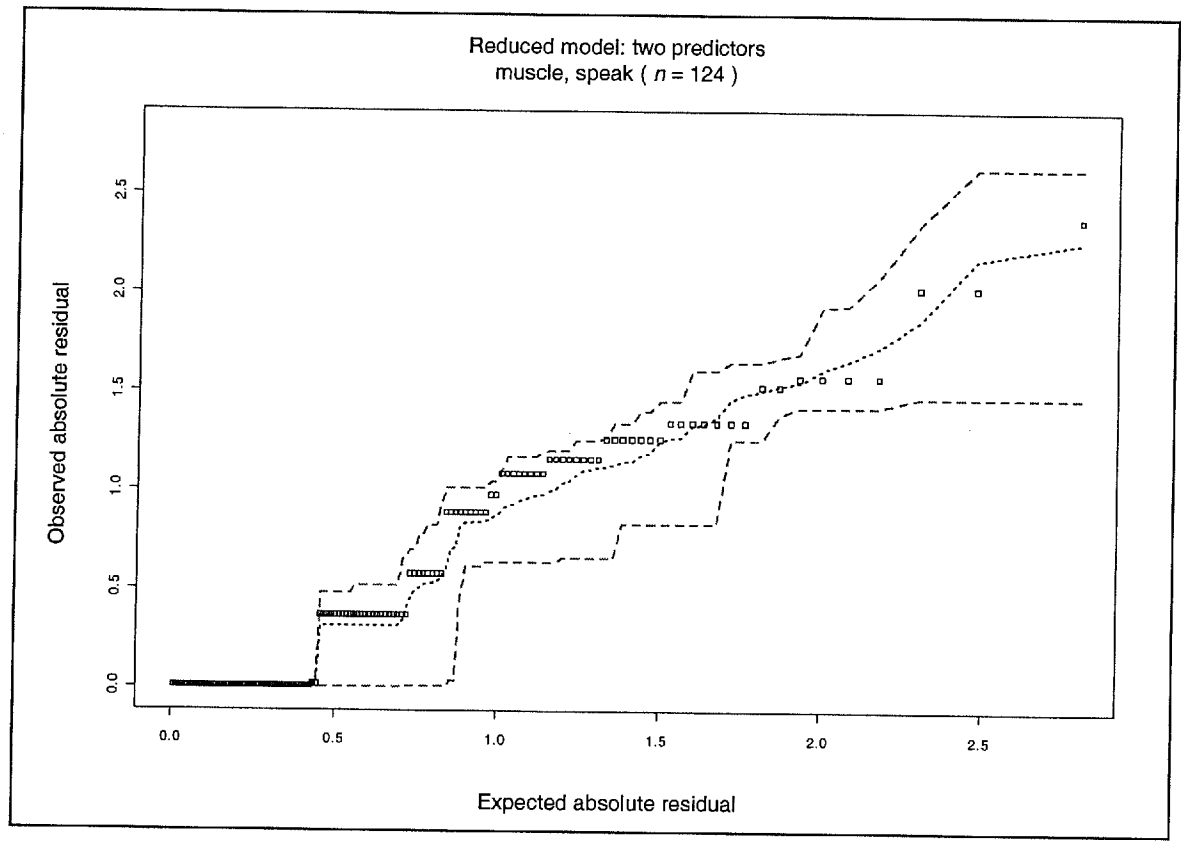


Figure 8

Half Normal Plot of Deviance Residuals with Simulated Envelope for Reduced Model with Two Predictors: Accessory Muscle Use, Ability to Speak



Comparing prevalence of severe distress as predicted versus as observed

Another way to assess how well the model explains the relationship between the predictors - in this case accessory muscle use, alterations in respiratory effort, and alterations in the ability to speak full sentences - and severe distress, is to assess how well the predicted prevalence of severe distress in the sample compares to what was actually observed. Tables 22 through 25 present these comparisons. Tables 22 and 23 contain comparisons for the model that retained accessory muscle use in the analysis, while Table 24 contains the comparison without accessory muscle use. Table 25 compares the model using only accessory muscle use and ability to speak full sentences. In these tables, each cell represents the prevalence of severe distress for the cases with that particular combination of findings. For example, in Table 22, for those cases that were not exhibiting accessory muscle use, the model predicted that 5.6% of those exhibiting mildly increased respiratory effort and a normal ability to speak would be in severe distress; 4.4% of those cases were actually observed to be in severe distress. Although not in perfect agreement, the estimates derived from the model without accessory muscle use more closely resemble those observed in the sample than the estimates derived from models that included accessory muscle use.

Tables 22 and 23

Comparison of Prevalence of Severe Distress between Predicted and Observed:

Accessory Muscle Use Retained in the Model

Table 22

Accessory Muscle Use Absent

Predicted				Observed			
Respiratory	**Ability to Speak**			**Respiratory**	**Ability to Speak**		
Effort	Normal	Bad	Worse	**Effort**	Normal	Bad	Worse
Normal	.000	.000	N/A[a]	Normal	.000	.000	N/A
Bad	.056	.196	.238	Bad	.044	.286	.500
Worse	.103	.318	.374	Worse	.000	.000	.600

[a] Indicates no cases in these cells

Table 23

Accessory Muscle Use Present

Predicted				Observed			
Respiratory	**Ability to Speak**			**Respiratory**	**Ability to Speak**		
Effort	Normal	Bad	Worse	**Effort**	Normal	Bad	Worse
Normal	.000	N/A[a]	N/A	Normal	.000	N/A[a]	N/A
Bad	.165	.446	.508	Bad	.000	.333	.500
Worse	.274	.607	.664	Worse	.667	.727	.500

[a] Indicates no cases in these cells

Tables 24

Comparison of Prevalence of Severe Distress between Predicted and Observed: Reduced Model with Ability to Speak and Respiratory Effort

Predicted				**Observed**			
Respiratory	**Ability to Speak**			**Respiratory**	**Ability to Speak**		
Effort	Normal	Bad	Worse	**Effort**	Normal	Bad	Worse
Normal	.000	.000	N/A[a]	Normal	.000	.000	N/A
Bad	.064	.312	.351	Bad	.042	.308	.500
Worse	.145	.529	.571	Worse	.200	.533	.533

[a] Indicates no cases in these cells

Table 25

Comparison of Prevalence of Severe Distress between Predicted and Observed: Reduced Model with Ability to Speak and Accessory Muscle Use

Predicted				**Observed**			
Accessory	**Ability to Speak**			**Accessory**	**Ability to Speak**		
muscle use	Normal	Bad	Worse	**muscle use**	Normal	Bad	Worse
Absent	0.033	0.204	0.292	Absent	.000	.000	N/A
Present	0.139	0.550	0.663	Present	0.400	0.588	0.500

[a] Indicates no cases in these cells

CHAPTER 5

DISCUSSION

The purpose of this study was to begin work on an Index of Respiratory Distress, which would prescribe a standard by which OOH providers could assess older patients in respiratory distress, and provide guidelines by which to identify those patients in severe distress. The specific aims were to 1) determine the appropriate items, scaling, and scoring of the instrument and 2) assess the instrument's reliability and validity. This chapter discusses the results in light of these aims. The chapter ends with a summary of the findings and a discussion of the limitations of the study; implications for practice, education, and research; and plans for future work in this area.

Determining the Appropriate Items, Scaling, and Scoring

Item Selection

Possible predictor variables were identified from the literature, from observing OOH providers in practice and learning about their assessment of patients in respiratory distress, and from reviewing OOH records to determine what the providers actually recorded. That list was refined for distribution to the EMS experts, and resulted in 25 items that represented all three domains in the definition of RD that was developed for this study, that is, problems with ventilation, oxygenation, or respiratory effort. Table 26 summarizes the domains represented by these potential predictors.

Table 26

Domains Represented by Potential Predictors Prior to Delphi Survey

Predictor	Domain
Nasal flaring	Effort[a], Ventilation[b]
Tracheal tugging	Effort, Ventilation
Retractions (intercostal, supraclavicular)	Effort
Accessory muscle use	Effort
Paradoxical respiratory movement	Ventilation
Position of patient in bed or chair	Effort, Ventilation
Altered blood pressure	Oxygenation
Altered respiratory rate	Ventilation, Oxygenation[c]
Altered heart rate	Oxygenation
Abnormal skin temperature	Oxygenation
Abnormal skin color	Oxygenation
Abnormal skin moisture	Oxygenation
Altered level or consciousness (LOC)/mentation	Oxygenation
Inability to speak full sentences	Effort
Pursed-lip breathing	Effort
Audible respiratory sounds	Ventilation
Dyspnea	Effort, Ventilation
Fatigue, "looking pooped"	Effort
Oxygen saturation (O2 sat)	Ventilation
Peak expiratory flow rate (PEFR)	Oxygenation
Respiratory effort	Effort
Facial expression of anxiety/stress/panic	Oxygenation
Agitation	Effort
Decreased exercise tolerance	Oxygenation
Pulsus paradoxus	Oxygenation

[a] Indicates problems with respiratory effort.
[b] Indicates problems with oxygenation
[c] Indicates problems with ventilation

Although placement of each predictor into the three domains may be open some different interpretations, the critical point is that all domains were represented. This list was shortened after the first Delphi survey, and the reduced list was then sent to the experts for ranking. The responses of the experts mirrored the lack of consensus noted in the literature review. For example, the two physicians agreed on the rank of only one predictor, and were as far apart as 9 points on one item. Similar inconsistencies were observed between the two nurses and the two paramedics. Considering the standard deviations as measures of the variability of rankings, some predictors had very little variability (e.g. abnormal skin temperature and abnormal skin moisture), and some showed a great deal of variability (e.g., altered blood pressure and altered respiratory rate). It is interesting to note that items with the lowest variability were items ***not*** included in the model. In fact, items ranked as the worst predictors had surprisingly little variability except for "altered blood pressure," and the wide standard deviation noted there can be accounted for by one rater who ranked it as 2; all other raters ranked it below 10. Although there was some consistency about what predictors probably weren't good indicators of severity, even the "experts" didn't agree about what were good indicators.

Choosing items for the final model was also complicated by the fact that many of the predictors were very similar, e.g., accessory muscle use and retractions. Some of the experts identified this problem and tried to remedy it in different ways, e.g. by arbitrarily assigning the same rank to more than one predictor, or eliminating predictors as redundant. The multiplicity of parameters used to assess the severity of respiratory

distress probably reflects a combination of individual differences in the ways patient respond to problems with ventilation, oxygenation, and effort, and the ways providers interpret these responses. The lack of consensus reflected in the experts' responses to the surveys provides additional evidence of the need to find some way to standardize both the assessment of these patients and the meanings of the findings.

The seven predictors identified for the model also represented all three domains included in the definition of respiratory distress. The domain of ventilation was represented by *altered respiratory rate* and *position of patient on arrival*; the domain of oxygenation was represented by *alterations in level of consciousness* and *decreased oxygen saturation levels*; and the remaining four predictors – *increased respiratory effort, the ability to speak full sentences*, and *accessory muscle use* are all in the domain of respiratory effort. Again, there may be arguments for placing these predictors in more than one domain, but what is important is that all domains were represented as the model was built.

Item Scaling and Scoring

Inasmuch as the scaling of the items on a 0-2 scale allowed the development of a robust model for assessing the association between the predictor score and the presence of severe distress, the scaling of the items was useful. However, developing the scoring rubric for the IRD depended on the results of both the logistic regression model and further analyses that were unable to be carried out due to the small sample size. Reasons for the smaller-than-anticipated sample size were discussed in part in the previous chapter

and additional discussion follows; the result was the investigator was unable to develop the scoring for the IRD.

Final Item Selection

A binary logistic regression model was built to identify the predictor variables most appropriate for inclusion in the IRD. The results indicated that accessory muscle use, inability to speak full sentences, and increased respiratory effort were the best predictors of severe respiratory distress in older OOH patients. However, there was substantial evidence in these results to indicate potentially serious threats to both their internal and external validity.

The extremely large standard errors in the parameter estimates for both levels of respiratory effort are the first indication of possible problems with these data. Although some of this may be a result of having no cases of severe distress with normal respiratory effort, this more probably represents the imprecision with which this predictor was measured. This is discussed below in more detail, but respiratory effort was the predictor that fit most poorly within the operational definition devised at the beginning of the study.

Additional evidence of potential problems with the model emerged as the investigator began exploring further reductions in the number of predictors. All of the models with only two predictors showed a significant decrease in explanatory power from the model with three predictors. Although the HL test showed no significant lack of fit in either of the two-predictor models, the DLOF test showed that removing respiratory

effort and leaving only ability to speak and accessory muscle use resulted in a model that fit poorly with the data. On the other hand, the model with only respiratory effort and ability to speak remaining showed the best agreement between the observed and predicted prevalence of severe respiratory distress for all combinations of predictors. These inconsistencies raise serious questions regarding the internal validity of the results.

There is also a clinical consideration that indicates a potential threat to external validity. The three predictors identified by the model, i.e., accessory muscle use, ability to speak full sentences, and respiratory effort, all reflect problems with ventilation and respiratory effort. This does not seem to be a clinically plausible result; patients with acute hypoxia usually reflect some alteration in cognition or behavior. Likewise, it is puzzling from a clinical standpoint to think that decreased oxygen saturation is not one of the predictors that was identified in the model. There are several possible explanation for this finding, the first of which is that the sample size may too small to detect meaningful variations in these predictors as they relate to the outcome. This may be related to another possible problem, that is, the levels of measurement may not be sensitive enough to identify these relationships. It may also be that OOH providers are not recording their findings consistently and accurately enough to detect meaningful differences. Of course, it is possible that the increased work of breathing manifested as labored breathing, accessory muscle use, and loss of ability to speak full sentences had succeeded in at least temporarily maintaining oxygenation status – and with it the level of consciousness. However, the investigator feels that these questionable results are most likely attributable

to the challenges she faced in defining and measuring both the predictor and outcome variables.

Limitations in defining and measuring the predictor variables

Stiell and his colleagues (Stiell & Wells, 1999) recommend that potential predictor variables in CDRs be clearly defined. This proved to be more difficult than anticipated. Although the investigator included a review of OOH medical records along with direct input from OOH providers in deriving the operational definitions for the items, when she began reviewing the OOH records, she had to interpret the entries far more extensively than planned. The problems with the operational definitions, along with the inconsistencies in charting, varied from predictor to predictor.

The problems encountered in the predictor *level of consciousness* revolved around the quality of the recordings rather than the quantity. All of the 311 charts reviewed contained an assessment of mental status, however, that assessment was most often limited to one of two very rough scales. The first was the Glasgow Coma Scale, which is not designed to evaluate subtle changes in mental status. The second was an assessment of four parameters that indicate gross level of consciousness, called the AVPU scale. In this scale, "A" indicates the patient is awake and alert; "V" indicates the patient responds to verbal stimulus only; "P" indicates the patient responds to painful stimulus only; and "U" indicates the patient is unconscious. There is no place in either the GCS or AVPU scale to note more subtle behavioral or cognitive changes, such as inability to follow directions or cooperate with treatment, that may be associated with acute hypoxemia. The

investigator strongly suspects that subtle changes such as those, although significant, went unrecognized or unrecorded by the OOH providers. It seems highly unlikely that only 3 patients among the 25 with severe respiratory distress exhibited some changes in mentation or behavior.

Recording the *ability to speak full sentences* was less consistent, and required greater interpretation on the part of the investigator. Almost 94% of original 311 charts contained some assessment of the patient's ability to speak full sentences, but this assessment did not always fall into the investigator's scaling plan. The goal was to differentiate those who could barely speak, from those whose speech was in some ways restricted, and that group from those who were speaking normally[6]. Problems in coding arose, for example, when OOH providers indicated that patients could speak, "2-4 word sentences." The investigator then had to decide if this belonged in category "0" (speaks 0-2 word sentences) or "1" (speaks 3-6 word sentences). Regardless of where the category limits were set, there was enough inconsistency in the OOH providers' assessments to necessitate a significant amount of interpretation on the part of the investigator.

The chief problem with the predictor *accessory muscle use* was the fact that it was recorded only 65% of the time. The investigator had planned to code accessory muscle use as not present if the providers did not mention it in their reports. However, this predictor was often omitted even for those patients whom the providers indicated *were* in

[6] Operational definitions for these categories were described in Table 12, above.

severe RD. Here the investigator was faced with the choice of coding the predictor as not present when she strongly suspected that was not so, or treating it as "missing." As a result, the investigator did not code accessory muscle use unless there was a statement in the record to indicate it either was or was not present. Since it appeared that the providers were recording this parameter on the sicker patients only, these results must be interpreted with great caution. It is especially interesting to note that the logistic regression model with only ability to speak and respiratory effort was the closest approximation to the observed data among all those tested; this also raises questions about the usefulness of accessory muscle use in predicting severe RD.

Out-of-hospital providers recorded data on the position of the patient upon their arrival on scene over 90% of the time. Here, again, however, interpreting these recordings was problematic. For example, if the medic wrote, "Arrived to find patient seated at the kitchen table," it is not clear whether that patient should be ranked as a 1, "found sitting up or high Fowler's," or 2, "found lying flat or sitting or reclining comfortably."

Information regarding *oxygen saturation* was present on almost 97% of all EMS records reviewed for this study, and very few fell outside the operational definitions used for assigning a category. Occasionally it was difficult to determine when supplemental oxygen had been administered in relation to the actual measurement of oxygen saturation, but this happened infrequently. The investigator is confident that these are reliable data.

Like oxygen saturation, information regarding *respiratory rate* was almost universally present on all 311 charts reviewed for this study (n=309; 99.6%), and the rates fit easily in the categories developed by the investigator.

Respiratory effort posed challenges in both the quality and quantity of the data recorded. Almost 31% of the 311 charts reviewed were missing data on respiratory effort, and few used precisely the language reflected in the operational definition. The operational definition was taken directly from a portion of the OOH record of a single EMS agency, and although the investigator was assured that all EMS provider agencies used essentially the same form for recording findings, this proved not to be the case. In addition, it turned out that only the first page of the two-page OOH record remained on the permanent hospital record, so that even for the 11 cases from this agency (8.8% of the sample), the data were not available in the precise form the investigator expected. Thus the original categories and operational definitions were essentially useless.

The goal in using three-tiered response categories for the predictors was to assign cases to one of three categories: normal, or no change from baseline (coded 2); not normal, but not extremely altered (coded 1); and showing extreme alteration from normal (coded 0). With that in mind, the investigator coded respiratory effort using the following general guidelines: if providers noted the patient to be in no distress, or it was noted that the patient was transported to the receiving facility "code 1," (no lights or sirens, no need to expedite transport), respiratory effort was coded as 2, or normal. If the providers noted patients were "in mild distress," "in some distress," "in moderate distress," or other

similar statement indicating some alteration from normal, respiratory effort was coded as 1, to indicate some, but not extreme, deviation from normal. If providers noted the patient was in "obvious" or "severe" respiratory distress, or that the patient exhibited labored respirations, respiratory effort was coded as 0.

The lack of a precise operational definition for respiratory effort is the most probable explanation for the very large standard errors associated with its parameter estimates, and the degree to which the investigator had to interpret the data in the OOH records makes these data highly suspect. Even though removing respiratory effort from the model significantly reduced its explanatory power, these results must be viewed with extreme caution.

This discussion would not be complete without considering the arbitrary nature of the limits that were set of each level within the predictor variables. As mentioned, the investigator assigned those limits based on her synthesis of the literature and her experience. As can be seen in the review of various asthma and croup CDRs included in Appendix 1, various investigators have used different boundaries in setting the levels of their predictor variables. Part of the ongoing work to develop the IRD will be to refine these limits to identify the findings that are the most sensitive predictors of severe distress.

Limitations in defining and measuring the outcome variable

Since there is no gold standard for "severe distress" against which to measure the performance of the predictors in the OOH setting, three proxy measures were chosen: a

statement by the receiving physician that the patient was in severe distress (or functional equivalent, e.g., "patient arrived *in extremis*"); the patient was placed on assisted ventilation (CPAP, BiPAP, or ETI) within one hour of arrival to the ED; or the patient was admitted to the ICU for treatment and monitoring of the respiratory problem. These measures were considered good approximations of the level of severity of the patient's respiratory distress in the field, even though they are assessed in the "Emergency Department Care" unit of service as depicted in the conceptual framework presented in Chapter 2.

Unlike the data for the predictor variables, data regarding these three proxies were invariably complete. Twenty-five cases in the analytic sample (20.2%) were designated as being in severe distress by the operational definition. This was similar to the prevalence in the larger sample of 311 ($\chi^2_{(1)} = .490$; p=.484). Since there is no agreed upon standard definition of severe distress, there is no way to assess whether this estimate is at all reflective of the prevalence of severe RD among all patients 50 years of age or older who access EMS for treatment and transport.

In addition to the issue of generalizability, there are clearly identifiable problems with the proxy variables. First, patients who were in severe distress in the OOH setting may improve en route to the ED and arrive in less than severe distress. By the same token, patients who were not in severe distress in the field may decompensate shortly after arrival to the ED. In this study, failure of the physician to record a general impression of the patient's respiratory status was interpreted as a "no," to the question,

"did the patient arrive in severe distress?" Aside from any possible problems with coding this proxy variable, there is no reason to believe that physicians are any more consistent in their assessments than are OOH providers. Finally, it is also possible – in fact, highly probable - that physicians' thresholds for placing patients on assisted ventilation or admitting them to the ICU will vary, raising questions regarding the reliability of these measures.

Despite these limitations, the investigator feels the choice of measures for the outcome variable is a sound one. There seems to be as great a likelihood that patients improved en route to the ED as that they decompensated shortly after arriving there. As for the variability in physician assessments and charting, the investigator believes this variability is likely distributed randomly in the sample. The sample was chosen from consecutive cases of patients who accessed EMS, and limited only by how complete their EMS documentation was; this would have no bearing on which physicians were represented in the sample.

The main advantage to using the three identified proxies is that it casts a relatively wide net on the sample. Ten patients who did not have statements in their chart indicating they arrived in severe distress were either placed on assisted ventilation or admitted to the ICU. Likewise, 15 patients who were not placed on assisted ventilation either arrived in severe distress or were admitted to ICU; and finally, eight patients admitted to ICU for their respiratory complaint were neither placed on assisted ventilation nor arrived in

severe distress. It is possible that these proxies overestimated the prevalence of severe distress; however, it is highly doubtful that they underestimated it.

It is interesting to note that the three predictors contained in the most parsimonious and best-fitting model are all predictors for which the investigator had difficulty either obtaining or coding the data. This may reflect a certain selection bias on the part of the OOH providers, in that they may have restricted their recording of findings on these variables largely to those with at least some level of respiratory distress. If this were true, and the investigator suspects that it is, it would be expected that these three variables are the best predictors of severe distress from among those in the model.

The issues raised in this discussion indicate significant threats to both the internal and external validity of these findings. As a result, the investigator views this study primarily as evidence that it is possible to identify the best predictors of severe distress in older out-of-hospital patients using clinical findings already being obtained by OOH providers. At this point, the investigator is not prepared to claim that she has identified the most appropriate items for the Index of Respiratory Distress; she is convinced, however, that by refining the operational definitions for the predictor variables and using a prospective design, this can ultimately be accomplished.

Other Challenges Encountered during the Study

Observing Out-of-Hospital Providers

Although the initial plan was for the investigator to observe a minimum of five contacts between OOH providers and patients in RD, she was unable to reach this goal

due to time constraints. This represents the well-known unpredictability of OOH practice; it is impossible to predict when providers will be called upon to care for any type of patient. Despite this, observations of providers assessing the respiratory status of patients who had other complaints, and the interviews with providers, resulted in rich descriptions of what providers included in their assessment and how they interpreted their findings. During the interviews, many OOH providers told the investigator they "just know" when patients are in severe respiratory distress; they respond to a total impression they have when they arrive to the scene. However, upon further questioning, providers began to list the clinical indicators they used and to describe how they interpreted the findings.

In addition to gathering the qualitative data for the first phase of this project, spending over 100 hours with OOH providers reminded the investigator of the constraints present in OOH practice. The confined spaces, ambient noise level, limited information, and time constraints clearly affected the providers' practice patterns e.g., patients were not undressed for assessments, lung sounds could not be assessed once the ambulance was moving, assessments and interventions were planned and executed to spend as little time as possible "on scene." One result of these constraints was that the investigator could not ride in the patient care compartment of the ambulance when family members were taken along with the patient, thus limiting observations. These limitations were partially offset by the willingness of OOH hospital providers to be "debriefed" after each patient care encounter.

Gaining Access to Out-of-Hospital Records

The processes of gaining access to appropriate medical records and then abstracting both the EMS and matching ED data were exercises in the vicissitudes of EMS research. Like Newgard & Lewis (2002), the investigator was caught in the changing atmosphere of OOH research when the OHSU IRB insisted on Federalwide Assurances (FWAs) for all EMS agencies from which the investigator planned to use patient-related data. Prior to this study, the IRB had apparently not required these assurances for OOH research with these agencies. After almost six months of work with local EMS agencies, only one of those that expressed interest in participating in the study actually succeeded in completing the FWA application process and obtaining the assurance. It took almost eight months to obtain initial IRB approval for the data abstraction phase of this study, and each modification of the protocol required from two to four weeks for approval. This was compounded by delays from the IRB at the receiving hospital from which most of the ED data were collected, and this hospital initially limited the investigator to 200 charts of patients. They were eventually persuaded to allow access to additional ED records, but by then new federal regulations designed to protect patient privacy had been implemented, and the requirements for review under these new guidelines added an additional four weeks to the approval process.

Access to the physical record (or electronic version thereof), as it turned out, was only one challenge. Equally critical was the limitation on the quality of the data in those EMS records. Some of this was alluded to in the discussions of each predictor variable,

but to emphasize, there was absolutely no consistency in what OOH providers recorded, where the findings were recorded, or how those findings related to the destination or OOH treatment decisions. The extent of these inconsistencies was far beyond what the investigator anticipated; in fact, only 19% of all the records originally identified as eligible for this study contained complete data on all seven predictors.

Although no formal analysis was done, the investigator did note occasional EMS chart entries of clinical indicators that were not included in the list of predictors for the IRD. Among these were *pursed-lip breathing, grunting respirations,* and *dyspnea.* In one instance the provider had a patient try to rate his/her level of dyspnea on a 1-10 scale. However, these were isolated instances; it did not appear that any significant number of providers was using these indicators in consistent manner.

Summary and Conclusions

The purpose of this study was to begin developing an Index of Respiratory Distress to assist out-of-hospital providers in caring for older patients in respiratory distress. The aims were to determine the items, scaling, and scoring for the IRD, and assess its reliability and validity. In retrospect, these were unrealistic goals given the time frame of the project and the special challenges faced by the investigator. Although falling short of the original aims of the project, some progress was made towards the overall goal, including: a) developing a concise definition of respiratory distress, b) assembling a thorough compilation of clinical indicators used by providers of all types to evaluate the severity of respiratory distress, c) using rigorous methods for reducing that list to the

indicators most likely to predict severe distress, and d) using multivariate statistics to evaluate how well those indicators differentiate older OOH patients in severe distress from those not in severe distress.

One of the most important findings of this study is the recognition that even with a relatively small sample size and imprecise measurements, three variables – accessory muscle use, ability to speak in full sentences, and respiratory effort - emerged as predictors of severe distress. The ability to identify severe respiratory distress in older OOH patients based on clinical predictors is critically important. Several studies have identified real and potential adverse effects of advanced life support interventions on these patients (Callaham, 1997; Pepe et al., 1994; Wuerz, Swope, Meador, Holliman, & Roth, 1994; Wuerz & Meador, 1992). This argues strongly for limiting the use of such therapies to the those who are in such severe distress that they cannot safely wait the 15 to 20 minutes it might take to get to the more controlled environment of the emergency department (Callaham, 1997). These results suggest it may be possible to make that determination by using data the OOH providers are already collecting. It will require much more work to identify the precise relationships between the most sensitive predictors and various levels of severity, but it would appear that such relationships are possible to detect.

Although the Emergency Medical Services Outcomes Project (EMSOP) identified respiratory distress as a priority for OOH outcomes research (Maio et al., 1999), they were unable to reach consensus on measures of severity of that distress (Maio, R.,

personal communication, 2001). Members of the EMSOP group identified many of the same challenges as did this investigator, namely the multiple diseases that manifest as respiratory distress, the variability in training among OOH providers, and the inter-agency and system differences in protocols and practices that result in manifestly different ways data are collected. This, in turn, makes it difficult to plan research in this area or generalize results from one setting to another. The EMSOP group's stated - although as yet unrealized - goal was to "identify measures that would not take much time, would not require exotic equipment or prolonged training, and could be performed by different levels of prehospital care providers" (Maio, personal communication, 2001). This study suggests looking more closely at some of the measures the OOH providers are already using and developing more precise ways of calibrating them and interpreting the findings.

The EMSOP group also suggested that measures of severity of respiratory distress can be used as both risk adjustment and outcome measures (Maio, R., personal communication, 2001). Here again, the results of this study concur with their work. Clinical indications of distress such as accessory muscle use, ability to speak full sentences, and respiratory effort may be measured by providers upon their arrival on scene and again upon arrival to the emergency department. Using Spaite, et al.'s (Spaite et al., 2001) conceptual model of OOH research, this would fit nicely in the *out-of-of hospital* unit of service for a single episode of care. Results of OOH and ED care could be risk-adjusted to the patients' initial level of severity, and the outcomes of various ALS

interventions could then be compared across different patient populations. All of this will have to await the results of the ongoing work on this project, but these initial findings suggest it may be possible to use clinical measures in this way.

Stiell & Wells (1999) identified six criteria for developing CDRs for use in emergency medicine. This study takes the process for developing the IRD through the first and into the second criteria. The needs for the IRD that have been identified include the prevalence of the problem and wide variations in practice among the OOH providers. The prevalence of the problem was identified clearly by the EMSOP group (Maio et al., 1999), and the wide variations in practice were manifest many places during this study. These became apparent right at the beginning with the lack of consensus among EMS experts and in the literature regarding the appropriate assessments and interpretation of findings for patients of any age in respiratory distress.

The variations in practice were especially apparent in the OOH records. For example, the investigator saw the *ability to speak 3-6 word sentences* described as evidence of three different levels of respiratory distress: mild, moderate, and severe. Similarly, some OOH providers listed several clinical findings such as accessory muscle use, position of patient, patient's ability to speak, etc., to support their assessment of level of severity while others simply described the patients as "mildly short of breath," "in some distress," or in "obvious respiratory distress," without providing any further information. It seems highly unlikely that the ability to speak 3-6 word sentences can

indicate three separate levels of distress; rather, these variations reflect the urgent need to standardize OOH practice in caring for older patients in respiratory distress.

This project is now focused on the second criterion described by Stiell & Wells (1999), defining and measuring both the predictor and outcome variables. Defining and measuring the predictors has proven especially challenging, as was noted above. The choice of clinical predictors was made through a combination of qualitative and quantitative research methods, and expert opinion was obtained using accepted methods for collecting and analyzing those opinions in rigorous ways. Clearly the operational definitions for predictor variables need to be refined and the measurement scales adjusted accordingly. On the other hand, the identification and measurement of the outcome variable was far easier and much less ambiguous. Although the investigator identified some potential sources of bias in the operational definition of severe distress, these were offset by the advantage of having an outcome variable that is both temporally close to OOH unit of service and plausibly related to the events that occurred during that time. In the continued absence of a widely accepted "gold standard," this definition seems to be adequate to the purpose.

The results of this study are subject to several limitations, many of which have already been described in detail. These include the difficulty in gaining access to OOH medical records, the variability in the quality of the data recorded on those records, the variations in practice among OOH providers, and problems with the operational definitions and coding of the predictor variables. These limitations resulted in a much

smaller sample size than originally planned, and as noted above, this made interpretation of the results of the logistic regression model somewhat problematic. Although it is premature to declare the three identified predictors as the "best" clinical predictors of severe respiratory distress in older OOH patients, it does seem reasonable to expect that once the limitations have been addressed, and a much larger sample can be obtained, the best clinical predictors can and will be identified.

It must also be emphasized that the results of this work, if they are generalizable at all, are limited to patients 50 years of age and older. The physical and behavioral responses of older adults experiencing problems in ventilation, oxygenation, and respiratory effort are different enough from those of other age groups (especially those of infants and children) to preclude applying findings from this study to any other group of patients. However, it is possible that the same methods used to develop the IRD may be used to develop a similar CDR for use with younger aged patients.

Completed development of the IRD would have important implications for clinical practice, for the education of OOH providers, and for OOH research. The implications for practice have been alluded to throughout the study, and include standardizing the assessment of older patients in respiratory distress, standardizing the interpretation of those findings, and providing guidelines for limiting advanced life support to those patients for whom the benefits outweigh the attendant risks of treatment. If accepted by the EMS community, this would also result in more consistent documentation across EMS agencies and regions.

The completed IRD would also streamline the education of OOH providers regarding the assessment and care of older patients in respiratory distress. Once the process of validating the IRD is complete, educators will be able to describe the relationship between selected clinical findings and the likelihood that patients are in severe distress. The IRD would probably be relied on most heavily by newer OOH providers who lack the clinical experience to "just know" when patients are in severe RD; more experienced providers can use it to record important clinical findings most succinctly.

Unlike implications for practice and education that need to await further development of the IRD, the implications of these findings for OOH research are immediate. First and foremost, these results imply that further research in this area is both important and possible to do, despite the challenges described in the study. Second, the findings imply that there is a relationship between clinical findings and the level of respiratory distress patients are experiencing. Third, this study demonstrates that, with perseverance, investigators can successfully navigate through the federal and local approval processes necessary to meet institutional requirements.

Plans for future work in this area include working with local EMS agencies to obtain the necessary federal certifications and access to their original OOH records. The investigator also plans to seek input from OOH providers in refining the operational definitions of the predictor variables. Once these two initial steps have been accomplished, the next step would be a prospective study in which OOH providers use an

assessment tool specifically designed to collect data on the predictor variables. The investigator would then obtain follow-up data at local receiving hospitals and proceed with the appropriate analysis.

Using prospectively collected OOH data would strengthen the findings of the ensuing multvariate analysis (Stiell & Wells, 1999). It would also allow an assessment of the inter-rater reliability of the assessment tool, since both OOH providers for each agency would be asked to complete the assessment independent of each other. It is anticipated that a multi-site study will ultimately be needed to obtain a large enough sample to generate meaningful results from multivariate statistical models and other appropriate statistical operations such as recursive partitioning and receiver operating curve analysis.

Despite the challenges described herein and the resulting limitations in this study, these preliminary results are promising. They represent a successful first effort to quantify an aspect of OOH care that has been neglected or avoided by other OOH researchers. With a more realistic appreciation of the amount of time it may take to develop it, the investigator is eager to proceed to the next step in developing a valid, reliable Index of Respiratory Distress.

REFERENCES

Abramson, N., & Safar, P. (1990). Deferred consent: Use in resuscitation research. *Annals of Emergency Medicine, 19*, 83-86.

Ailani, R. K., Ravakhah, K., DiGiovine, B., Jacobsen, G., Tun, T., Epstein, D., & West, B. C. (1999). Dyspnea differentiation index. *Chest, 116*, 1100-1104.

Altman, D. G., & Royston, P. (2000). What do we mean by validating a prognostic model? *Statistics in Medicine, 19*, 453-473.

Baxt, W. G., Berry, C. C., Epperson, M. D., & Scalzitti, V. (1989). The failure of prehospital prediction rules to classify trauma patients accurately. *Annals of Emergency Medicine, 18*, 21-28.

Bestall, J. C., Paul, E. A., Garrod, R., Garnham, R., Jones, P. W., & Wedzicha, J. A. (1999). Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax, 54*, 581-586.

Bledsoe, B. E., Porter, R. S., & Shade, B. (1994). *Paramedic Emergency Care* (Second ed.). Englewood Cliffs, NJ: Brady Prentice Hall.

Brain Trauma Foundation. (2000). *Management and Prognosis of Severe Traumatic Brain Injury* (Clinical Guidelines): Brain Trauma Foundation and American College of Neurological Surgeons, Joint Section on Neurotrauma and Critical Care.

Brice, J. H., Garrison, H. G., & Evans, A. T. (2000). Study design and outcomes in out-of-hospital emergency medicine research: A ten-year analysis. *Prehospital Emergency Care, 4*, 144-150.

Cales, R. H. (1986). Injury severity determination: Requirements, approaches, and applications. *Annals of Emergency Medicine, 15*, 1427-1433.

Callaham, M. (1997). Quantifying the scanty science of prehospital emergency care. *Annals of Emergency Medicine, 30*, 785-790.

Callaham, M., & Madsen, C. D. (1996). Relationship of timeliness of paramedic advanced life support interventions to outcomes in out-of-hospital cardiac arrest treated by first responders with defibrillators. *Annals of Emergency Medicine, 27*, 638-648.

Callaham, M. L. (1991). *Current Practice of Emergency Medicine*. Philadelphia: B.C. Decker.

Caroline, N. (1991). *Emergency Medical Treatment: A text for EMT-As and EMT - Intermediates* (Third ed.). Boston: Little Brown and Compnay.

Caroline, N. (1995). *Emergency Care in the Streets* (Fifth ed.). Boston: Little, Brown, and Company.

Collett, D. (1991). *Modelling Binary Data*. London: Chapman & Hall.

Cook, S., & Meek, P. M. (2001, April 19-21, 2001). *Self-management of pulmonary symptoms: How much is too much?* Paper presented at the Health Care Challenges Beyond 2001: Mapping the Journey for Research and Practice, Seattle, WA.

Cordell, W. H., Olinger, M. L., Kozak, P. A., & Nyhuis, A. W. (1994). Does anybody really know what time it is? Does anybody really care? *Annals of Emergency Medicine, 23*, 1032-1036.

Craven, R. F., & Hirnle, C. J. (2000). *Fundamentals of Nursing: Human Health and Function* (Third ed.). Philadelphia: Lippincott.

Cronin, S. N. (1997). Respiratory disorders. In J. M. Black & E. Matassarin-Jacobs (Eds.), *Medical-Surgical Nursing: Clinical Management for Continuity of Care* (Fifth ed., pp. 1105-1131). Philadelphia: W.B. Saunders.

Crosby, L. A., & Lewallen, D. G. (Eds.). (1997). *Emergency Care and Transportation of the Sick and Injured* (Sixth, Revised ed.). Rosemont, IL: American Academy of Orthopedic Surgeons.

Cummins, R. O. (1999). Why are researchers and emergency medical services managers *not* using the Utstein Guidelines? *Academic Emergency Medicine, 6*, 871-875.

Dantzker, D. R., MacIntyre, N. R., & Bakow, E. D. (1995). *Comprehensive Respiratory Care*. Philadelphia: W.B. Saunders Company.

Davies, P. J., & Hoffman, L. (2000). Nursing Management of Respiratory Failure. In S. M. Lewis & M. M. Heitkemper & S. R. Dirksen (Eds.), *Medical Surgical Nursing: Assessment and Management* (Fifth ed., pp. 1895-1913). St. Louis: Mosby.

Delbridge, T., Domeier, R., & Key, C. B. (2003). Prehospital asthma management. *Prehospital Emergency Care, 7*, 42-47.

Des Jardins, T. (1990). *Clinical Manifestations of Respiratory Disease* (Second ed.). Chicago: Year Book Medical Publishers, Inc.

Dettenmeier, P. A. (1996). *Pulmonary Nursing Care*. St. Louis: Mosby.

Eckstein, M., Chan, L., Schneir, A., & Palmer, R. (2000). Effect of prehospital advanced life support on outcomes of major trauma patients. *The Journal of Traum: Injury, Infection, and Critical Care, 48*, 643-648.

Eisner, M. D., Katz, P. P., Yelin, E. H., Henke, J., Smith, S., & Blanc, P. D. (1998). Assessment of asthma severity in adults with asthma treated by family practitioners, allergists, and pulmonologists. *Medical Care, 36*, 1567-1577.

Emerman, C. L., Shade, B., & Kubincanek, J. (1991). A comparison of EMT judgment and prehospital triage instruments. *The Journal of Trauma, 31*, 1369-1375.

Fishman, A. P. (Ed.). (1998). *Fishman's Pulmonary Diseases and Disorders* (Third ed.). New York: McGraw-Hill.

Garrison, H. G., Maio, R. F., Spaite, D. W., Desmond, J. S., Gregor, M. A., O'Malley, P. J., Stiell, I. G., Cayten, C. G., Chew, J. L., MacKenzie, E. J., & Miller, D. R. (2002). Emergency Medical Services Outcomes Project III (EMSOP III): The role of risk adjustment in out-of-hospital outcomes research. *Annals of Emergency Medicine, 40*, 79-88.

Gausche, M., Lewis, R. J., Stratton, S., Haynes, B. E., Gunter, C., Goodrich, S., Poore, P., McCollough, M., Henderson, D., Pratt, F., & Seidel, J. (2000). Effect of out-of-

hospital pediatric endotracheal intubation on survival and neurological outcome: A controlled clinical trial. *JAMA, 283*, 783-790.

George, R. B., Light, R. w., Matthay, M. A., & Matthay, R. A. (1995). *Chest Medicine: Essentials of Pulmonary and Critical Care Medicine*. Baltimore: Williams and Wilkins.

Gibson, G. (1973). Evaluative criteria for emergency ambulance systems. *Social Science and Medicine, 7*, 425-454.

Grant, H. D., Murray, R. H. J., & Bergeron, J. D. (1994). *Emergency Care* (Sixth ed.). Englewood Cliffs NJ: Brady Prentice Hall.

Hafen, B. Q., Karren, K., & Mistovich, J. J. (1996). *Prehospital Emergency Care* (Fifth ed.). Upper Saddle River NJ: Brady Prentice Hall.

Hajiro, T., Nishimura, K., Tsukino, M., Ikeda, A., Koyama, H., & Izumi, T. (1998). Analysis of clinical methods used to evaluate dyspnea in patients with chronic obstructive pulmonary disease. *American Journal of Respiratory Care Medicine, 158*, 1185-1189.

Hoekstra, J. W., Banks, J. R., Martin, D. R., Cummins, R. O., Pepe, P. E., Stueven, H. A., Jastremski, M., Gonzalez, E., & Brown, C. G. (1993). The effect of first-responder automated defibrillation on time to therapeutic interventions during out-of-hospital cardiac arrest. *Annals of Emergency Medicine, 22*, 1247-1253.

Holleran, R. S. (1994). *Prehospital Nursing*. St. Louis: Mosby.

Howard, J. M. (2000). Historical background to "Accidental Death and Disability: The Neglected Disease of Modern Society." *Prehospital Emergency Care, 4*, 285-289.

Howell, J. M. (Ed.). (1998). *Emergency Medicine* (Vol. I). Philadelphia: W.B. Saunders.

Ihaka, R., & Gentleman, R. (2003). R: A Language and Environment (Version 1.7.1): The R Development Core Team.

Jacobs, S., Shortland, G., Warner, J., Dearden, A., Gataure, P. S., & Tarpey, J. (1994). Validation of a croup score and its use in triaging children with croup. *Anaesthesia, 49*, 903-906.

Janson-Bjerklie, S., Ferketich, S., Benner, P., & Becker, G. (1992). Clinical markers of asthma severity and risk. *Heart & Lung, 21*, 265-272.

Karras, D. J., Sammon, M. E., Terregino, C. A., Lopez, B. L., Griswold, S. K., & Arnold, G. K. (2000). Clinically meaningful changes in quantitative measures of asthma severity. *Academic Emergency Medicine, 7*, 327-334.

Kendrick, K. R. (2000). Can a self-rating 0-10 scale for dyspnea yield a common language that is understood by ED nurses, patients, and their families? *Journal of Emergency Nursing, 26*(3), 233-234.

Laupacis, A., Sekar, N., & Stiell, I. G. (1997). Clinical prediction rules: A review and suggested modifications of methodological standards. *JAMA, 277*, 488-494.

Lett, R. R., Hanley, J. A., & Smith, J. S. (1995). The comparison of injury severity instrument performance using likelihood ratio and ROC curve analyses. *The Journal of Trauma: Injury, Infection, and Critical Care, 38*, 142-148.

Levitsky, M. D. (1999). *Pulmonary Physiology* (Fifth ed.). New York: McGraw Hill.

Liberman, M., Mulder, D., & Sampalis, J. (2000). Advanced or basic life support for trauma: A meta-analysis and critical review of the literature. *The Journal of Trauma: Injury, Infection, and Critical Care, 49*, 585-599.

Mahler, D. A., Jones, P. W., & Guyatt, G. H. (1998). Clinical measurement of dyspnea. In D. A. Mahler (Ed.), *Dyspnea* (Vol. 111, pp. 149-199). New York: Marcel Dekker.

Maio, R. F., Garrison, H. G., Spaite, D. W., Gregor, M. A., Cayten, C. G., Chew, J. L., Hill, E. M., Joyce, S. M., MacKenzie, E. J., Miller, D. R., O'Malley, P. J., & Stiell, I. G. (1999). Emergency Medical Services Outcomes Project I (EMSOP I): Prioritizing conditions for outcomes research. *Annals of Emergency Medicine, 33*, 423-432.

Maslanka, A. M. (1993). Scoring systems and triage from the field. *Emergency Clinics of North America, 11*, 15-27.

McGinn, T. G., Guyatt, G. H., Wyer, P. C., Naylor, C. D., Stiell, I. G., & Richardson, W. S. (2000). Users' guides to the medical literature XXII: How to use articles about clinical decision rules. *JAMA, 284*, 79-84.

Mosesso, V. N. (1993). The most neglected tool in EMS: The clock. *Annals of Emergency Medicine, 22*, 1311-1312.

Mosesso, V. N., Dunford, J., Blackwell, T., & Griswell, J. K. (2003). Prehospital therapy for acute congestive heart failure: State of the art. *Prehospital Emergency Care, 7*, 13-23.

Moy, M. L., Lantin, M. L., Harver, A., & Schwartzstein, R. M. (1998). Language of dyspnea in assessment of patients with acute asthma treated with nebulized albuterol. *American Journal of Respiratory Care, 158*, 749-753.

Narad, R. A., & Driesbock, K. R. (1999). Regulation of ambulance response times in California. *Prehospital Emergency Care, 3*, 131-135.

National Heart, Lung, and Blood Institute. (1997). *Guidelines for the Diagnosis and Management of Asthma* (Expert Panel Report: Clinical Practice Guidelines NIH No. 97-4051): National Institutes of Health National Heart, Lung, and Blood Institute.

National Highway Traffic Safety Administration (1996). *Emergency Medical Services Agenda for the Future* (DOT HS 808 441). Washington, DC: U.S. Government Printing Office.

Newberry, L. (Ed.). (1998). *Sheehy's Emergency Nursing: Principles and Practice* (Fourth ed.). St. Louis: Mosby.

Newgard, C. D., & Lewis, R. J. (2002). The paradox of human subjects protection in research: some thoughts on and experiences with the Federalwide Assurance program. *Academic Emergency Medicine, 9*, 1426-1429.

Newgard, C. D., Lewis, R. J., & Jolly, B. T. (2002). Use of out-of-hospital variables to predict severity of injury in pediatric patients involved in motor vehicle crashes. *Annals of Emergency Medicine, 39*, 481-491.

Nichol, G., Detsky, A. S., Stiell, I. G., O'Rourke, K., Wells, G., & Laupacis, A. (1996). Effectiveness of emergency medical services for victims of out-of-hospital cardiac arrest: a metaanalysis. *Annals of Emergency Medicine, 27*, 700-709.

Nichol, G., Stiell, I. G., Blackburn, J., Luciano, T., Nesbitt, L., Wells, G. A., & Huszti, E. (2003). Quality of life outcomes for respiratory distress patients treated by EMS. *Academic Emergency Medicine, 10*(5), 449.

Nicholl, J., & Willoughby, J. (1998). Effects of ambulance response times are being evaluated. *British Medical Journal, 317*, 684.

Ornato, J. P., Doctor, M. L., Harbour, L. F., Peberdy, M. A., Overton, J., Racht, E. M., Zauhar, W. G., Smith, A. P., & Ryan, K. A. (1998). Synchronization of timepieces to the atomic clock in an urban emergency medical services system. *Annals of Emergency Medicine, 31*, 483-487.

Parkin, P. C., Macarthur, C., Saunders, N. R., Diamond, S. A., & Winders, P. M. (1996). Development of a clinical asthma score for use in hospitalized children between 1 and 5 years of age. *Journal of Clinical Epidemiology, 49*, 821-825.

Parshall, M. B. (1999). *Exacerbated Chronic Dyspnea in Emergency Patients.* Unpublished Doctoral dissertation, University of Kentucky.

Parshall, M. B., Welsh, D., Brockopp, D. Y., Heiser, R. M., Schooler, M. P., & Cassidy, K. B. (2001). Reliability and validity of dyspnea sensory quality descriptors in heart failure patients treated in an emergency department. *Heart and Lung, 30*, 57-65.

Pepe, P. E. (1993). Out-of-hospital resuscitation research: Rationale and strategies for controlled clinical trials. *Annals of Emergency Medicine, 22*, 25-30.

Pepe, P. R., Abramson, N. S., & Brown, C. G. (1994). ACLS-Does it really work? *Annals of Emergency Medicine, 23*, 1037-1041.

Polit, D. F., & Hungler, B. P. (1999). *Nursing Research: Principles and Methods* (Sixth ed.). Philadelphia: Lippincott.

Reinke, L. F., & Hoffman, L. A. (2000). Nursing Assessment: Respiratory System. In S. M. Lewis & M. M. Heitkemper & S. R. Dirksen (Eds.), *Medical Surgical Nursing: Assessment and Management* (Fifth ed., pp. 553-578). St. Louis: Mosby.

Rose, L. (1980). A policy perspective on research in Emergency Medical Services Systems. *Emergency Health Services Quarterly, 1*(1), 99-110.

Rosen, P. (Ed.). (1998). *Emergency Medicine: Concepts and Clinical Practices* (Fourth ed. Vol. II). St. Louis: Mosby.

Schuh, S., Johnson, D., Stephens, D., Callahan, S., & Canny, G. (1997). Hospitalization patterns in severe acute asthma in children. *Pediatric Pulmonology, 23*, 184-192.

Senkowski, C. K., & McKenney, M. G. (1999). Trauma scoring systems: A review. *Journal of the American College of Surgeons, 189*, 491-503.

Smith, S. R., Baty, J. D., & Hodge, D. (2002). Validation of a pulmonary score: An asthma severity score for children. *Academic Emergency Medicine, 9*, 99-104.

Spaite, D. W., Criss, E., Valenzuela, T. D., & Guisto, J. (1995). Emergency medical services systems research: Problems of the past, challenges of the future. *Annals of Emergency Medicine, 26*, 146-151.

Spaite, D. W., Criss, E., Valenzuela, T. d., & Meislin, H. W. (1997). Developing a foundation for the evaluation of expanded-scope EMS: A window of opportunity that cannot be ignored. *Annals of Emergency Medicine, 30*, 791-796.

Spaite, D. W., Criss, E. A., Valenzuela, T. D., & Meislin, H. W. (1998). Prehospital advanced life support for major trauma: Critical need for clinical trials. *Annals of Emergency Medicine, 32*, 480-489.

Spaite, D. W., Maio, R., Garrison, H. G., Desmond, J. S., Gregor, M. A., Stiell, I. G., Cayten, C. G., Chew, J. L., MacKenzie, E. J., Miller, D. R., & O'Malley, P. J. (2001). Emergency Medical ServicesOutcomes Project (EMSOP) II: Developing the foundations and conceptual models for out-of-hospital research. *Annals of Emergency Medicine, 37*, 657-663.

Spaite, D. W., Valenzuela, T. D., Meislin, H. W., Criss, E. A., & Hinsberg, P. (1993). Prospective validation of a new model for evaluating emergency medical services systems by in-field observation of specific time intervals in prehospital care. *Annals of Emergency Medicine, 22*, 638-645.

SPSS for Windows. (Standard Version)(2000). [Statistics package].

Steele, D. W., Santucci, K. A., Wright, R. O., Natarajan, R., McQuillen, K. K., & Jay, G. D. (1997). Pulsus Paradoxus: An objective measure of severity in croup. *American Journal of Critical Care Medicine, 156*, 331-334.

Stiell, I. G., DeMaio, V. J., Nesbitt, L., Nichol, G., Brisson, D., & Beaudoin, T. (2003). Predictors of survival for out-of-hospital respiratory distress patients in the OPALS study. *Academic Emergency Medicine, 10*, 431.

Stiell, I. G., Wells, G. A., Spaite, D. W., Nesbitt, L., Cousineau, D., DeMaio, V. J., Campeau, T., Dagnone, E., Nichol, G., Field, B. J., Beaudoin, T., & Brisson, D. (2003). OPALS Study Phase III: What is the impact of Advanced Life Support on out-of-hospital cardiac arrest? *Academic Emergency Medicine, 10*, 423.

Stiell, I. G., & Wells, G. A. (1999). Methodologic standards for the development of clinical decision rules in emergency medicine. *Annals of Emergency Medicine, 33*, 437-447.

Stiell, I. G., Wells, G. A., DeMaio, V. J., Spaite, D. W., Field, B. J., Munkley, D. P., Lyver, M. B., Luinstra, L. G., & Ward, R. (1999). Modifiable factors associated with improved cardiac arrest survival in a multicenter basic life support/defibrillation system: OPALS study phase I results. *Annals of Emergency Medicine, 31*, 44-50.

Stiell, I. G., Wells, G. A., Field, B. J., Spaite, D. W., DeMaio, V. J., Ward, R., Munkley, D. P., Lyver, M. B., Luinstra, L. G., Campeau, T., Maloney, J., & Dagnone, E. (1999). Improved out-of-hospital cardiac arrest survival through the inexpensive

optimization of an existing defibrillation program: OPALS study phase II. *JAMA, 281*, 1175-1181.

Stiell, I. G., Wells, G. A., Spaite, D. W., Graham, N., Nesbitt, L., DeMaio, V. J., Lyver, M. B., Brisson, D., Martin, M. T., Doherty, J., Tammy, & Cousineau, D. (2002). Multicenter controlled clinical trial to evaluate the impact of advanced life support on out-of-hospital respiratory distress patients. *Academic Emergency Medicine, 9*, 357.

Stiell, I. G., Wells, G. A., Spaite, D. W., Nesbitt, L., Cousineau, D., DeMaio, V. J., Campeau, T., Dagnone, E., Nichol, G., Field, B. J., Beaudoin, T., & Brisson, D. (2003). OPALS Study Phase III: What is the impact of Advanced Life Support on out-of-hospital cardiac arrest? *Academic Emergency Medicine, 10*, 423.

Streiner, D. L., & Norman, G. R. (1995). *Health Measurement Scales: A Practical Guide to Their Use and Development* (Second ed.). Oxford: Oxford University Press.

Swor, R. A. (1999). Out-of-hospital cardiac arrest and the Utstein style: Meeting the customer's needs? *Academic Emergency Medicine, 6*, 875-877.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (Fourth ed.). Boston: Allyn and Bacon.

Taussig, L. M., Castro, O., Beaudry, P. H., Fox, W. W., & Bureau, M. (1975). Treatment of laryngotracheobronchitis (croup). *American Journal of Diseases of Children, 129*, 790-793.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. *The Lancet* (July 13, 1974), 81-84.

Thelan, L. A., Lough, M. E., Urden, L. D., & Stacy, K. M. (1998). *Critical Care Nursing: Diagnosis and Management* (Third ed.). St. Louis: Mosby.

Thompson, J. M., McFarland, G. K., Hirsch, J. E., & Tucker, S. M. (1997). *Mosby's Clinical Nursing* (Fourth ed.). St. Louis: Mosby.

Traver, G., Mitchell, J. T., & Flodquist-Priestly, G. (1991). *Respiratory Care: A Clinical Approach.* Gathersburg: Aspen.

Wasson, J. H., Sox, H. C., Neff, R. K., & Goldman, L. (1985). Clinical prediction rules: Applications and methodological standards. *New England Journal of Medicine, 313*, 793-799.

West, J. G., Murdock, M. A., Baldwin, L. C., & Whalen, E. (1986). A method for evaluating field triage criteria. *The Journal of Trauma, 26*, 655-659.

White, E. B. (1997). Forward. In R. L. Kane (Ed.), *Understanding Health Care Outcomes Research* (pp. 265). Gaithersburg: Aspen Publishers, Inc.

Wilson, R. C., & Jones, P. W. (1989). A comparison of the visual analogue scale and modified Borg scale for the measurement of dyspnoea during exercise. *Clinical Science, 76*, 277-282.

Wilson, S. F., & Thompson, J. M. (1990). *Respiratory Distress.* St. Louis: Mosby.

Wuerz, R., Swope, G., Meador, S., Holliman, J. C., & Roth, G. S. (1994). Safety of prehospital nitroglycerine. *Annals of Emergency Medicine, 23*, 31-36.

Wuerz, R. C., & Meador, S. A. (1992). Effects of prehospital medications on mortality and length of stay in congestive heart failure. *Annals of Emergency Medicine, 21*, 669-674.

Appendix 1: Asthma and Croup Scoring Systems

Author/Yr*	Type of System	Pt. Population	Purpose	Description	Results	Comments
Jacobs, et al. 1993	Croup	Pediatric, 3 mos. to 9.5 years old	Validate a simple scoring system upon which to base triage for admission to pediatric ICU	Four part instrument evaluating stridor, cyanosis, sternal retractions, respiratory rate (based on weight), pulse rate (based on age). Score ranged 0-3 on some items, 0-2 on others. Higher scores indicate more severe signs/symptoms. Scores >5 resulted in admission to ICU	Good results for sensitivity, specificity, predictive values, and accuracy on both phases of study.	Well-done study with good model for testing instrument. Instrument itself is too complex for use in prehospital. Some components may transfer well to adult population.
National Heart, Lung, and Blood Institute (NHLBI), 1997	Asthma	Adults and children	To be used in the absences of peak expiratory flow measurements to help patients assess the severity of their exacerbations	A large chart displaying signs, symptoms, and functional assessments, and providing examples of findings when patients are having mild, moderate, or severe exacerbations. Also includes parameters for "respiratory arrest imminent." No scores attached to assessments.	No testing was reported in the guidelines.	While this may provide a useful beginning for tool development, it must be verified with clinical data from prehospital records and providers

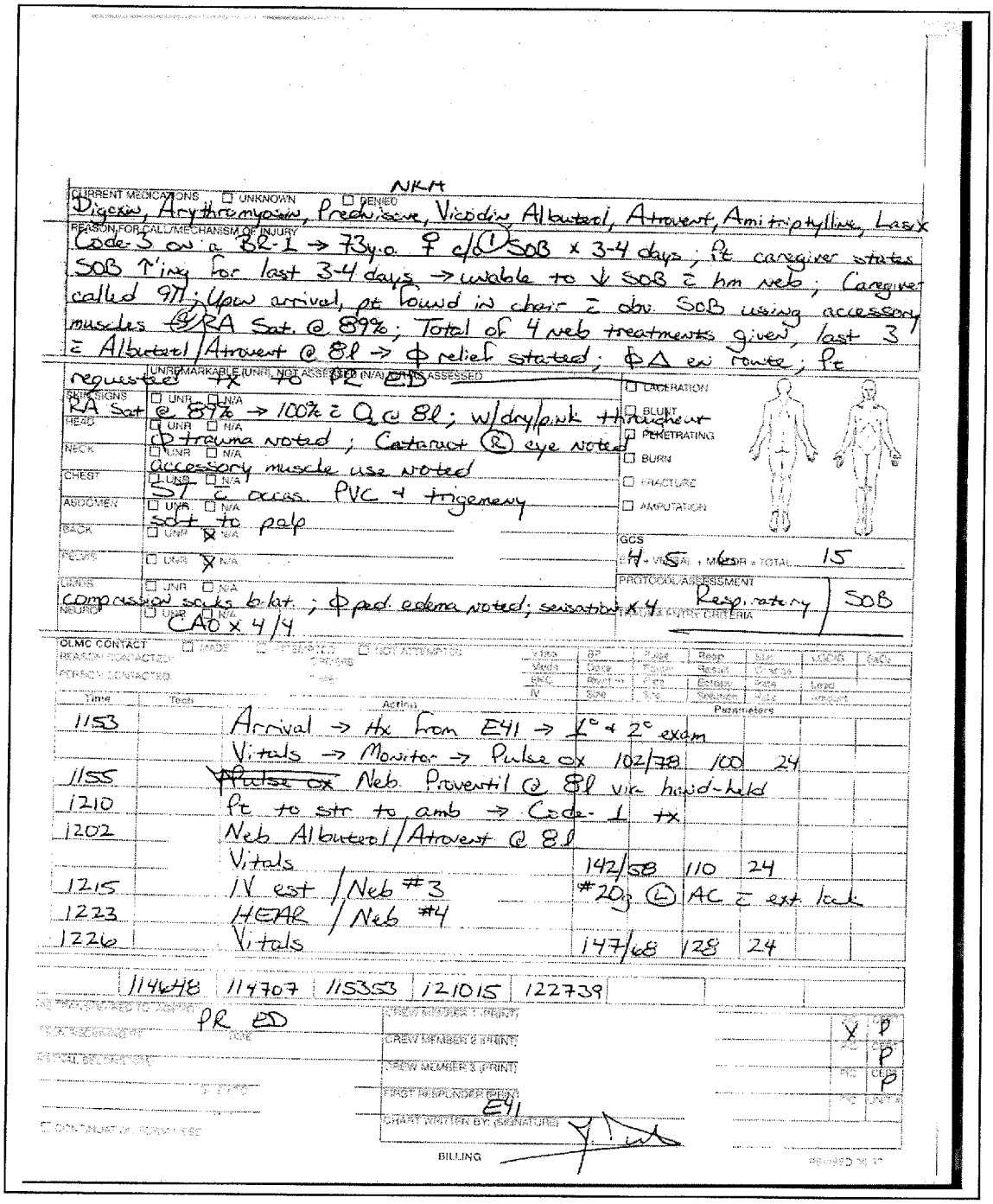
* All studies appear in the reference list beginning on page 142.

Author/Yr*	Type of System	Pt. Population	Purpose	Description	Results	Comments
Parkin, et.al. 1996	Asthma	Pediatric, 1-5 years old	To develop and evaluate asthma score in hospitalized children for use to evaluate responses to therapy.	Five part instrument. Evaluates respiratory rate, wheezing(heard with stethoscope), indrawing (retractions), observed dyspnea (impression of the observer of the degree of patient's breathlessness), and inspiratory/expiratory ratio. Possible score 0-2 on each part. Highest possible score = 10. Higher score indicates more severe signs/symptoms	High inter-rater reliability; good discriminatory power; good face and construct validity; adequately sensitive	Items culled from literature. Inspiratory/expiratory ratio had poor inter-rater reliability, but added to discriminatory power. While report indicated instrument to be valid and reliable, it is too cumbersome for prehospital use.
Schuh, et al. 1997	Asthma	Children, 5-17 years old	To predict hospitalizatio n for asthmatics in the ED	Three part instrument. Evaluated accessory muscle use, wheezing, and "dyspnea," (defined as activity level and ability to speak full sentences). Each part was rated on a 0-3 point scale, and then totaled. Highest possible score = 9. The higher the score, the more severe the signs/symptoms	Instrument was found useful in predicting the probability of hospitalization after 2 hours of ED treatment when combined with a peak expiratory flow rate (PEFR) of less than 50% of predicted	Developed locally for use in one study. There was no report of reliability or validity of the instrument. Reliance on PEFR makes it inappropriate for prehospital use.
Smith, et al., 2002	Asthma	Children 5-17 years old	To validate the pulmonary	A three-item instrument, with scores ranging form	PS correlated well with PEFR	In use at that facility for many years. No

Author/Yr*	Type of System	Pt. Population	Purpose	Description	Results	Comments
			score (PS) as a measure of airway obstruction	0-3 on each item (higher scores indicate greater obstruction. Items are: respiratory rate (by age), wheezing, and accessory muscle use (sternocleidomastoid only).	as measure of obstructions, i.e., the higher the PS (indicating worse obstruction) the lower the PEFR.	information as to how it is used; no predictive data supplied. Use only in children in ED.
Steele, et al. 1997	Croup	Pediatric, 2.5-14 years old	To evaluate the use of *pulsus paradoxus* as an objective measure of severity of upper airway obstruction	Measured by a finger arterial pressure probe	Correlated well with a clinical croup score.	Did not report of validation done with croup score used as criterion measure. Equipment used entirely unsuitable for prehospital use.
Taussig, et al., 1975	Croup	Pediatric, 5 mos. to 11 years old	Used to determine patients' responses to different treatment regimens	Five-part instrument. Parts include color, air entry, retractions, level of consciousness, and stridor. Range of scores 0-3 for 4 items, 0-2 for level of consciousness. Highest possible score = 14. Higher score indicates more severe signs/symptoms.	Authors found instrument was not sensitive enough to changes in conditions in patients with less than moderate to severe symptoms	One of the earliest instruments reported in literature. Instrument developed locally; not validated.

Appendix 2: Sample Out-of-Hospital Medical Record

Side 1



NKA

CURRENT MEDICATIONS ☐ UNKNOWN ☐ DENIED
Digoxin, Erythromycin, Prednisone, Vicodin, Albuterol, Atrovent, Amitriptyline, Lasix

REASON FOR CALL/MECHANISM OF INJURY
Code 3 on a BLS-1 → 73y.o. ♀ c/o SOB x 3-4 days; Pt caregiver states SOB ↑'ing for last 3-4 days → unable to ↓ SOB ∑ hm neb; Caregiver called 911; Upon arrival, pt found in chair ∑ obv. SOB using accessory muscles @ RA Sat. @ 89%; Total of 4 neb treatments given, last 3 ∑ Albuterol/Atrovent @ 8l → ϕ relief stated; ϕΔ en route; Pt requested tx PR ED

UNREMARKABLE (UNR) NOT ASSESSED (N/A) ABNORMAL ASSESSED

SKIN SIGNS ☐ UNR ☐ N/A
RA Sat @ 87% → 100% ∑ O2 @ 8l; w/dry/pink thready skin

HEAD ☐ UNR ☐ N/A
ϕ trauma noted; Cataract (R) eye noted

NECK ☐ UNR ☐ N/A
accessory muscle use noted

CHEST ☐ UNR ☐ N/A
ST ∑ occas. PVC + trigemeny

ABDOMEN ☐ UNR ☐ N/A
soft to palp

BACK ☐ UNR ☒ N/A

PELVIS ☐ UNR ☒ N/A

LIMBS ☐ UNR ☐ N/A
compression socks b/lat.; ϕ ped. edema noted; sensation x4

NEURO ☐ UNR ☐ N/A
CAO x 4/4

☐ LACERATION ☐ BLUNT ☐ PENETRATING ☐ BURN ☐ FRACTURE ☐ AMPUTATION

GCS E 4 + V 5 + M 6 = TOTAL 15

PROTOCOL/ASSESSMENT
Resp. Emergency | SOB

OLMC CONTACT ☐ MADE ☐ ATTEMPTED ☐ NOT ATTEMPTED
REASON FOR CONTACT
PERSON CONTACTED

Time	Teen	Action	Parameters						
1153		Arrival → Hx from E41 → 1° + 2° exam							
		Vitals → Monitor → Pulse ox	102/78	100	24				
1155		~~Pulse ox~~ Neb Proventil @ 8l via hand-held							
1210		Pt to str to amb → Code-1 tx							
1202		Neb Albuterol/Atrovent @ 8l							
		Vitals	142/58	110	24				
1215		IV est / Neb #3	#20g (L) AC ∑ est lock						
1223		HEAR / Neb #4							
1226		Vitals	147/68	128	24				

114648	114707	115353	121015	122739

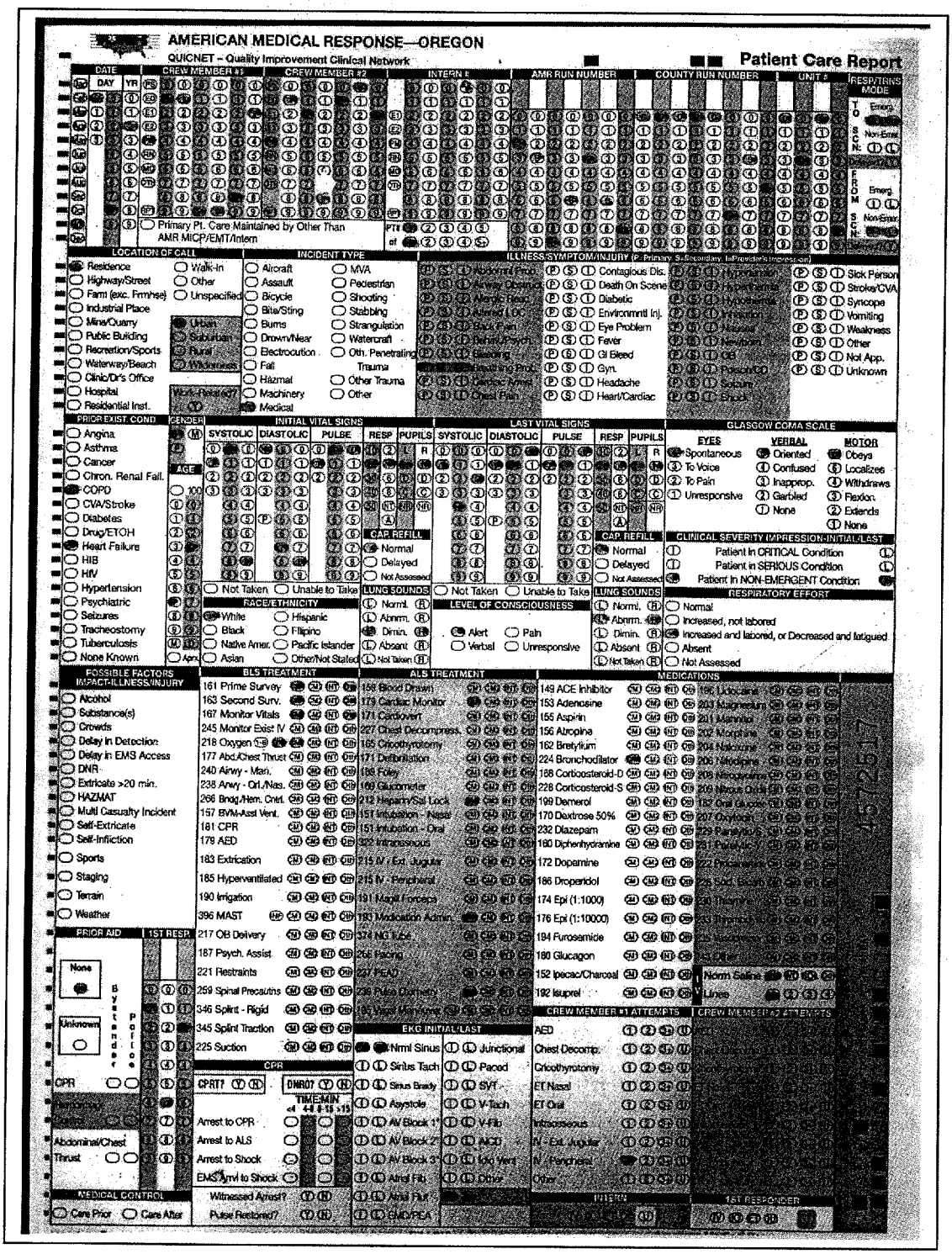
PR ED

CREW MEMBER 1 (PRINT) X P
CREW MEMBER 2 (PRINT) P
CREW MEMBER 3 (PRINT) P
FIRST RESPONDER (PRINT) E41
CHART WRITTEN BY: (SIGNATURE)

BILLING

Appendix 2: Sample Out-of-Hospital Medical Record

Side 2



Appendix 3: Approved EMS Consent

IRB# 7113-1
Approved: 08-06-2002

OREGON HEALTH & SCIENCE UNIVERSITY
Informed Consent Form: Out-of-Hospital Provider

TITLE: The Index of Respiratory Distress: A Clinical Decision Rule to Assist Out-of-Hospital Providers in Caring for Older Patients in Respiratory Distress

PRINCIPAL INVESTIGATOR: Susan E. Shapiro, RN, MS, CEN

DISSERTATION ADVISOR: Linda McCauley, Ph.D., RN, FAAN

PURPOSE:

You have been invited to participate in this research study because you regularly assess and treat out-of-hospital patients with complaints of respiratory distress or problems breathing. The purpose of this study is to develop a tool that will assist providers such as you in assessing the severity of the distress these patients are experiencing.

Your participation in this study will last up to 12 hours or one shift.

It is anticipated that between 8 and 15 additional providers will be invited to participate as well.

PROCEDURES:

Your participation in this study will involve being observed by the investigator while you engage in the routine care and treatment of patients experiencing respiratory problems. The investigator will not participate in any way, but may take notes of her observations.

After your patient has been delivered to the emergency department, the investigator will ask you a series of questions regarding what she observed during the patient encounter.

RISKS AND DISCOMFORTS

You will not experience any risks associated with this study. Although you will not experience any physical discomforts, you may experience minor psychological discomfort at the idea of being observed or while answering questions. These questions will focus only on how you assessed your patient and how you arrived at your treatment decision.

In addition, you may find it somewhat inconvenient having an additional person on the scene of a call or riding in the ambulance with you. It is not expected that the presence of the investigator will interfere in any way with your ability to provide your usual level of care to patients. However, if you find that the observation is hindering you in any way, you may ask to have the observation stopped.

BENEFITS:

You may or may not personally benefit from participating in this study. However, by serving as a participant, you may contribute new information that may benefit patients in the future.

ALTERNATIVES:

You may choose not to participate in this study. Doing so will in no way affect your relationship with the investigator, OHSU, or your employing agency.

CONFIDENTIALITY:

Neither your name nor your identity will be used for publication or publicity purposes.

Information collected during the observation and interview will be identified by a unique identifier. The investigator and her dissertation advisor will have access to any records linking your name with that identifier, and this information will be destroyed as soon as the observation and interviews are completed. The OHSU Institutional Review Board may review and photocopy research-related records. No information linking your name to this study will be entered into a computer.

Your employer will not have access to any information linking you to the study or its outcome.

COSTS:

There are no costs associated with your participation in this study.

LIABILITY:

The Oregon Health & Science University is subject to the Oregon Tort Claims Act (ORS 30.260 through 30.330). If you suffer any injury and damage from this research project through the fault of the University, its officers or employees, you have the right to bring legal action against the University to recover the damage done to you subject to the limitations and conditions of the Oregon Tort Claims Act. You have not waived your legal rights by signing this form. For clarification on this subject, or if you have further questions, please call the OHSU Research Support Office at (503) 494-7887.

PARTICIPATION:

Susan E. Shapiro (503) 494-3865, or 503-452-8009, has offered to answer any other questions you may have about this study. If you have any questions regarding your rights as a research subject, you may contact the OHSU Research Support Office at (503) 494-7887.

You may refuse to participate, or you may withdraw from this study at any time without affecting your relationship with or treatment at the Oregon Health & Science University. If you choose to withdraw, the information obtained from you up to that point will not be included in the study.

You will be given a copy of this consent for your records.

SIGNATURES:

Your signature below indicates that you have read the foregoing and agree to participate in this study.

_____ _____

Signature of Subject Date

_____ _____

Susan E. Shapiro, RN, MS, CEN Date
Principal Investigator

Appendix 4: OOH Provider Observation and Debriefing Guidelines

Date & (day of week): **Time of encounter:** **Time of interview:**

Level of provider: (highest level of training/certification/licensure): [] EMT-Basic [] EMT-Intermediate [] EMT paramedic [] RN [] Other (specify)

Age: **Race:** (investigator-identified): **# years in OOH practice**

Encounter data: Patient age **Patient race**

Level of distress per OOH provider **per SES**

What did provider actually assess?

What did provider say he/she used/relied on to determine severity of patient's RD?

Appendix 5: Guidelines for Out-of-Hospital Provider Interviews

(This interview may be conducted with one or more providers at the same time, after the patient has been delivered to the emergency department)

Introduction:

I am interested in learning how you arrived at the decision to treat (or not treat) the patient you just took to the emergency department. What was it that convinced you that this patient was (really sick) (not all that sick)?

Possible probes:

1. What role did vital signs, i.e., pulse rate, depth, quality and respiratory rate, depth, and quality, play in your decision?

2. What role did the patient's level of consciousness, mentation (thinking process), and behavior play in your decision?

3. What role did skin signs (temperature, color, moisture) play in your decision?

4. How did you decide what the patient's level of severity was?

Other:

Are there any assessment parameters you prefer over others to indicate severity? If so, why?

Are there any other things you saw, heard, or otherwise sensed that guided your decision making?

Is there anything else about your decision making you think I should know?

Appendix 6: Initial Survey for Out-of-Hospital Experts

School of Nursing
Oregon Health & Science University
3181 SW Sam Jackson Park Road
Portland, OR 97201
shapiros@ohsu.edu
(Spaite et al., 1993)

Dear Colleague,

Thank you for agreeing to assist me in the development of the Index of Respiratory Distress. As was explained in our phone call, I am requesting your expert opinion on which items to include in this clinical decision tool. This packet contains a place for you to describe your involvement in OOH care and three rating forms. The rating forms present various assessment parameters out-of-hospital (OOH) providers might use to assess the severity of respiratory distress in their patients. One form asks your opinion about the importance of the parameter as an indicator of illness severity; the second asks the degree to which you think OOH providers can evaluate this parameter in a rapid and consistent manner; and the third asks how likely you think the providers are to assess this parameter routinely in their practice.

These parameters were identified through literature review, chart review, observation of providers during ride-alongs, and debriefing interviews with participating providers after caring for patients in respiratory distress. If, however, you feel an important parameter has been omitted from this list, please let me know in the space provided at the end of the materials.

When you have completed these enclosures, please send them back to me in the stamped, self-addressed envelope provided in the packet. As we discussed, after the analysis of this information is complete, you will be contacted again to review the final list of parameters being considered.

Thank you again for your willingness to assist with this project; your input is critical to its success. If you have any questions, please feel free to call me at 503-452-8009, or use the other contact information noted above.

Sincerely,

Susan E. Shapiro, RN MS, CEN
Doctoral candidate
Oregon Health & Science University

Summary of Your Out-of-Hospital Experience

Level of Provider (medic, nurse, or physician):

Current Position:

Years in current position:

Years involved in OOH care:

1. As an OOH provider (EMT or Paramedic);_____
2. As a nurse _____
3. As a physician_____
4. As a multi-licensed or certified provider (please specify which licenses/certifications) _____

In summary, in what ways have you been involved in OOH care in the past five years?

Have you taught OOH care to physicians, nurses, and/or EMTs/Paramedics? If so, please summarize this experience.

Have you done OOH research? If so, please summarize this experience?

The Importance of the Parameter as an Indicator of the Severity of Disease

Instructions: Please review the parameters below, and rate them in terms of how reflective of severity of disease you think abnormal findings are when seen in older patients in respiratory distress. I am not looking for parameters you think make the dianosis easiest; rather, *I am interested in those you think indicate the greatest degree of physioloigical compromise or threat to life.*

1=not reflective of disease severity; 2=probably a good indicator of disease severity; 3= a critical indicator of disease severity

Parameter			
Nasal flaring			
Tracheal tugging			
Retractions			
Accessory muscle use			
Paradoxical respiratory movement			
Position of patient when provider arrives			
Altered blood pressure			
Altered respiratory rate			
Altered heart rate			
Abnormal skin temperature			
Abnormal skin color			
Abnormal skin moisture			
Level or consciousness/mentation			
Inability to speak full sentences			
Pursed-lip breathing			
Audible respiratory sounds			
Dyspnea			
Fatigue			
Decreased oxygen saturation			
Compromised peak expiratory flow rate			
Increased respiratory effort			
Facial expression indicating anxiety/stress/panic			
Agitation			
Decreased exercise tolerance			
Pulsus paradoxis			

The Ability of Providers to Assess Parameter in Rapid, Consistent Manner

Instructions: Please review the assessment parameters below and rate them in terms of how likely you think OOH providers are to be able to assess them in older patients in a rapid, consistent manner in the OOH environment (either at the scene or en route). Another way of thinking about this is to think about which parameters you think OOH providers assess most reliably.

1=not likely to be assessed consistently or reliably; 2=can be assessed consistently or reliably; 3= very likely to be assessed consistently or reliably

Parameter			
Nasal flaring			
Tracheal tugging			
Retractions			
Accessory muscle use			
Paradoxical respiratory movement			
Position of patient when provider arrives			
Altered blood pressure			
Altered respiratory rate			
Altered heart rate			
Abnormal skin temperature			
Abnormal skin color			
Abnormal skin moisture			
Level or consciousness/mentation			
Inability to speak full sentences			
Pursed-lip breathing			
Audible respiratory sounds			
Dyspnea			
Fatigue			
Decreased oxygen saturation			
Compromised peak expiratory flow rate			
Increased respiratory effort			
Facial expression indicating anxiety/stress/panic			
Agitation			
Decreased exercise tolerance			
Pulsus paradoxis			

How Likely are OOH Providers to Assess These Parameters Routinely

Instructions: Please review the assessment parameters below and rate them in terms of which ones you think OOH providers use routinely in assessing their older patients in respiratory distress. A rating of 1 indicates you think OOH providers rarely assess this parameter, while a rating of 3 indicates you think OOH providers assess this routinely on almost every call of this nature. I recognize that some parameters such as heart tate and respiratory rate are done on all patients all the time and this will be considered in how these results are analyzed.

1=rarely if ever assessed for; 2=assessed for by some medics some of the time; 3= almost always assessed for

Parameter			
Nasal flaring			
Tracheal tugging			
Retractions			
Accessory muscle use			
Paradoxical respiratory movement			
Position of patient when provider arrives			
Altered blood pressure			
Altered respiratory rate			
Altered heart rate			
Abnormal skin temperature			
Abnormal skin color			
Abnormal skin moisture			
Level or consciousness/mentation			
Inability to speak full sentences			
Pursed-lip breathing			
Audible respiratory sounds			
Dyspnea			
Fatigue			
Decreased oxygen saturation			
Compromised peak expiratory flow rate			
Increased respiratory effort			
Facial expression indicating anxiety/stress/panic			
Agitation			
Decreased exercise tolerance			
Pulsus paradoxis			

What Have I Left Out?

Thank you, again, for taking the time to complete this form. If you think I have omitted an important assessment parameter that should be included in this clinical decision tool, please provide that for me in the space below. Your assistance with this project is greatly appreciated.

Other parameters that should be considered:

Other comments or suggestions regarding this project:

Appendix 7: Second Survey for Out-of-Hospital Experts

School of Nursing
Oregon Health & Science University
3181 SW Sam Jackson Park Road
Portland, OR 97201
shapiros@ohsu.edu

January 11, 2003

Dear :

Thank you again for helping me with this project. After reviewing the responses from your first efforts, I have compiled a list of the fifteen assessment parameters you all identified as being the most important indicators of physiological compromise in older patients in respiratory distress, and most likely to be reliably assessed by out-of-hospital providers, and most likely to be incorporated into those providers' practices. I have two requests of you this time. First, I'd like you to ***rank order*** these variables from 1 (the best predictor) to 15 (the worst predictor) on the list. You may use whatever criteria you choose for establishing the rank. I would also like you to review the list and indicate two things: a) whether there is any item on the list you think shouldn't be there, and b) whether there is something not on the list you think absolutely should be there.

I appreciate you taking the time to complete this short questionnaire. Any input you provide will assist tremendously in making this a sound clinical decision rule.

I have included a stamped, self-addressed envelope for your convenience. If you are interested in the outcome of this ranking exercise, please let me know and will forward the results to you.

As always, please contact me if you have any questions or concerns.

Sincerely,

Susan E. Shapiro, RN MS, CEN
Doctoral candidate
Oregon Health & Science University
503-452-8009

Instructions: Please review the list of assessment parameters provided below. These have been identified by you and your peers as being the most important indicators of physiological compromise in patients in respiratory distress, and most likely to be reliably assessed by out-of-hospital providers, and most likely to be incorporated into those providers' practices. Now I'd like you to ***rank order them from 1-15***, with 1 indicating the best predictor and 15 indicating the worst. ***Please do not rank any two items at the same level.***

In addition, I'd like you to indicate if you feel there's an item on this list that shouldn't belong there, and if so, why, using the key provided on the next page.

Finally, please add to the list any predictor(s) you think should be on this list but isn't.

Predictor	Rank	**Reason for Removing (if applicable)**
Retractions		
Accessory muscle use		
Position of patient when provider arrives –relaxed vs.semi-fowlers vs. high fowlers/tripod		
Altered blood pressure		
Altered respiratory rate		
Altered heart rate		
Abnormal skin temperature		
Abnormal skin moisture		
Level of consciousness/mentation		
Inability to speak full sentences		
Pursed-lip breathing		
Audible respiratory sounds		
Dyspnea – new or significantly increased from baseline		
Decreased oxygen saturation		
Increased respiratory effort		
Add other items here		

Key for reasons for removing any predictors:

a. This is not an important indicator of physiological compromise.
b. Out-of-hospital providers cannot assess this parameter consistently or reliably.
c. Out-of-hospital providers would not do this routinely
    c.1 It takes too long
    c.2 Out-of-hospital providers cannot assess this reliably
    c.3 It is inappropriate for OOH providers to do this in the field
d. Other. Please specify:

Appendix 8: Out-of-Hospital Data Abstraction Form

Data Abstraction Form: OOH Records

Study ID number_____

Date of call_____ Time of call_____

Levels of providers on call [] ALS []BLS

Reason for call/patient's chief complaint_____

Patient age_____ Patient race_____

Predictors:

Level of consciousness		0=unconsious, 1=mild alteration, 2=normal
Ability to speak		0=unable –2 words, 1= 3-5 word sentences, 2=normal or not noted
Accessory muscle use		0=present, 2=none noted
Position of patient		0= tripod, upright and leaning forward, 1= semi-fowlers/head elevated, 2=flat or not remarkable
Oxygen saturation		0=<95% on 100%, 1=≥95% with O2, 2=≥95% on room air
Respiratory rate		0=Respiratory rate < 6 or > 32 per minute, 1= between 6-8 or 20-32 per minute, 2= between 8 and 20 per minute
Respiratory effort		0= Rate increased and labored, or patient apneic, 1= Rate increased but not labored, 2= Normal

Provider impression (field diagnosis) []COPD [] asthma []CHF [] pneumonia [] other _____

ALS field treatment, if any:

[] oxygen [] IV [] albuterol/ipatropium [] morphine [] furosemide

[] nitroglycerine [] other _____

OOH provider's assessment of clinical severity: 1 (critical); 2 (serious); 3 (non-emergent)

Appendix 9: Emergency Department Data Abstraction Form

Data Abstraction Form: ED Record

Study ID number_____

Date and time of arrival to ED_____
Discharge date & time_____

ED discharge diagnosis 1_____

2 _____

3 _____

1. Statement that patient arrived in severe distress [] yes [] no

2. Patient was intubated/put on CPAP within first hour in the ED [] yes [] no

3. Patient was admitted to an intensive care unit [] yes [] no

4. Comments

Appendix 10: Out-of-Hospital Provider Participants

Provider	**Agency Type**	**Level of Practice**	**Years' Experience**
A1	Transport	Paramedic	10
A2	Transport	Paramedic	<1
A3	Transport	Paramedic	6
A4	Transport	Paramedic	2
A5	Transport	Paramedic	12
T1	Fire	Paramedic	19
T2	Fire	Paramedic	3
T3	Fire	Basic EMT	12
T4	Fire	Paramedic	10
P1	Fire	Paramedic	6
P2	Fire	Paramedic	27
P3	Fire	Paramedic	13
P4	Fire	Paramedic	7

Appendix 11: Summary of ALS Contacts

Contact Number	OOH Provider Level	OOH Years Experience	Patient Age	Level of Distress per OOH Provider	Level of Distress per Investigator	Observed Assessments Done by OOH Provider	OOH Provider Report on Important Assessment Parameter
1	P	10	78	None	Mild	General appearance, SaO_2, pulse rate (via SaO_2,) lung sounds	"She was just sitting there in a chair, talking in full sentences"
2	P	19	68	Mild	Mild	SaO_2, mentation, skin color, vital signs, lung sounds	Dyspnea, skin color, mentation, patient's ability to speak easily
3	P	6	90	"didn't look good"	Moderate, in general. Resp. distress mild	SaO_2, general appearance, vital signs	Saw patient lying still, pale, not breathing hard. "She didn't look like she was having a breathing problem."
4	P	13	78	None	None	General appearance, lung sounds, SaO_2	Observed patient "from 5 feet away...saw her face - she wasn't in any distress." Face, "not anxious or fearful." She was "not breathing hard; no

Contact Number	OOH Provider Level	OOH Years Experience	Patient Age	Level of Distress per OOH Provider	Level of Distress per Investigator	Observed Assessments Done by OOH Provider	OOH Provider Report on Important Assessment Parameter
							accessory muscle use or nasal flaring." Re: lung auscultation: "wanted to be sure she hadn't aspirated or anything and wasn't junky, but she was clear." "She was satting at 98% before she got any oxygen, so I knew she was okay."
5	P	10	34	None	None	Vital signs, SaO_2, skin signs	The fact that the patient was lying flat in bed, respiratory rate grossly normal; skin color good, warm, and dry

P = paramedic
SaO_2 = oxygen saturation level

Appendix 12: Experts' OOH Professional Involvement

Profession	Current Position	Years in Current Position	Years in OOH Work	OOH involvement	OOH Teaching	OOH Research
M1- MD	Faculty, Emergency Medicine	13	24	On-line medical direction and research	Physical assessment, ACLS, research techniques, various disease/injury-related topics	Epidemiologist, methodologist, QA, trauma care, outcomes research
M2- MD	EMS Medical Director	13	20	Medical director, protocol development, ride-alongs, discipline	Teaching medics, writing book chapters and articles	30+ (yrs?), mostly observational studies (published) and numerous unpublished studies
R1-RN	Clinical manager, emergency services	<1	14	Paramedic 14 years, ED nurse 10 years, clinical instructor for paramedics, EMT students rotate through ED	Adjunct clinical faculty for EMT/Paramedic programs, presentations at emergency nursing functions re: interfacing with EMS, ED residents ride with me on the ambulance, ACLS/ PALS/ BCLS/ to EMTs, paramedics, nurses, and MDs using OOH situations	None
R2 – RN	Emergency and Trauma	4	2.5	Flight nurse in system that did	Have taught in paramedic programs in the past and	none

Profession	Current Position	Years in Current Position	Years in OOH Work	OOH involvement	OOH Teaching	OOH Research
	CNS			primary scene response; paramedic 2.5 years. Currently doing volunteer paramedic work at local FD	currently for local FD	
P1-Paramedic	Captain/ paramedic	5	20	Field paramedic; community college and university instructor; collegiate curriculum development committees; private ambulance paramedic; fire service EMS instructor; private industry medical/safety consultant	Community college instructor; paramedic training institute instructor; university adjunct faculty-clinical instructor; ACLS instructor; BLS instructor trainer	QA activities in 2 fire departments for 15 years, e.g. Utstein model cardiac arrest registry; trending EMS sentinel and frequency performance indicators; National EMS benchmarking project; paramedic ECG interpretation and treatment correlation study.
P2-Paramedic	Firefighter paramedic	4	11	Paramedic, county EMS; firefighter paramedic (non-transporting paramedic ALS first responder in	One-two hour classes on various EMS subjects, including respiratory distress to EMTs and First Responders.	None

Profession	Current Position	Years in Current Position	Years in OOH Work	OOH involvement	OOH Teaching	OOH Research
				combination urban, suburban, rural settings).		

Appendix 13: Initial Tabulation of Expert's Score from First Survey

Predictor Variable	A: Total Severity Score	B:Total Reliability Score	C: Total Usability Score	Total A +B	Total A+B+C
LOC/mentation	17	17	17	34	51
Respiratory rate	15	18	18	33	51
Inability to speak full sentences	15	18	15	33	48
Pursed lip breathing	15	14	11	29	40
Retractions	15	13	13	28	41
Decrease SaO2	14	18	18	32	50
Position upon arrival	14	14	16	28	44
Accessory muscle use	14	14	13	28	41
Skin temp	14	14	13	28	41
Dyspnea	14	12	15	26	41
Compromised PEFR	14	9	7	23	30
Increased respiratory effort	13	13	15	26	41
Fatigue	13	10	10	23	33
Paradoxical resp. movement	13	8	8	21	29
Audible resp. sounds	12	16	16	28	44
Skin moisture	12	14	16	26	42
Facial expression of anxiety/stress/panic	12	12	12	24	36
Nasal flaring	12	12	11	24	35
Pulsus Paradoxus	12	8	6	20	26
Pulse rate	11	18	18	29	47
Agitation	11	13	14	24	38
Tracheal tugging	11	9	8	20	28
B/P	9	17	18	26	44
Skin color	9	13	17	22	39
Decreased exercise tolerance	9	6	6	15	21
Median scores	**13**	**13**	**14**	**26**	**41**
Mean	**12.8**	**13.2**	**13.24**	**26**	**39.24**
Standard deviation	**2.06**	**3.46**	**3.94**	**4.56**	**7.94**

Appendix 14: Results of Second Survey of EMS Experts

Predictor	M1	M2	R1	R2	P1	P2	Total	Median Rank	Mean score	SD	# raters placing in top 5	# raters placing in top 10
LOC/mentation	1	1	2	8	2	8	22	2	3.67	3.39	4	6
Inability to speak full sentences	2	5	11	1	3	1	23	2.5	3.83	3.82	5	5
Position of patient	3	4	1	6	8	2	24	3.5	4.00	2.61	4	6
Decreased SaO2	6	3	10	4	4	3	30	4	5.00	2.68	4	6
Accessory muscle use	7	7	5	2	9	4	34	6	5.67	2.50	2	6
Dyspnea	5	12	8	9	6	5	45	7	7.50	2.74	2	5
Increased resp. effort	15	6	8	5	7	7	48	7	8.00	3.58	1	6
Pursed-lip breathing	12	9	4	7	11	6	49	8	8.17	3.06	1	4
Altered Resp. rate	4	10	8	10	1	11	44	9	7.33	3.98	2	5
Retractions	10	8	6	3	10	12	49	9	8.17	3.25	1	4
Audible respiratory sounds	13	11	3	11	5	9	52	10	8.67	3.88	2	3
Abnormal skin moisture	9	13	12	12	12	10	68	12	11.33	1.51	0	1
Altered BP	11	2	13	15	15	14	70	13.5	11.67	4.97	1	1
Altered heart rate	8	14	9	13	14	15	73	13.5	12.17	2.93	0	0
Abnormal skin temperature	14	15	14	14	13	13	83	14	13.83	0.75	0	
Median							**48**					
Mean							**47.6**					
Standard deviation							**19.22**					

Appendix 15: Results of Logistic Regression with All Seven Predictors in Model

Coefficient	Parameter Estimate	Standard Error	Wald Statistic	*p value*
(intercept)	-24.692	320.414	-0.07	.939
LOC[1] 1[2]	-1.297	1.386	-0.94	.350
LOC 0[3]	-0.325	1.344	-0.242	.809
Accessory Muscle[4]	1.117	0.667	1.674	.094
Oxygen[5] 1	-0.252	1.233	-0.204	.838
Oxygen 0	-0.272	1.326	-0.205	.838
Position[6] 1	-0.084	0.950	-.088	.930
Position 0	0.408	0.970	0.421	.674
Effort[7] 1	12.254	201.295	0.061	.952
Effort 0	12.864	201.295	0.064	.949
Rate[8] 1	9.914	249.315	0.040	.968
Rate 0	9.857	249.315	0.040	.969
Speak[9] 1	1.368	0.821	1.666	.096
Speak 0	1.782	0.969	1.840	.066

[1] LOC= level of consciousness
[2] 1 = mild alteration from normal or baseline
[3] 0 = significant alteration from normal or baseline
[4] Accessory muscle = accessory muscle use present
[5] Oxygen = oxygen saturation
[6] Position = position of patient when OOH provider arrived
[7] Effort = respiratory effort
[8] Rate = respiratory rate
[9] Speak = ability to speak in full sentences