Development and Validation of a Data Abstracting Tool for the Evaluation of Quality of Care in Emergency Medical Services

By Douglas Lorne Andrusiek

B.A. (Hons), The University of Western Ontario, 1987

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In

THE FACULTY OF GRADUATE STUDIES
(Health Care and Epidemiology)

THE UNIVERSITY OF BRITISH COLUMBIA

July 2005

© Douglas Lorne Andrusiek, 2005
Abstract

INTRODUCTION: Peer review is common in Emergency Medical Services (EMS); however its reliability and validity is not well studied. We sought to determine the reliability and criterion validity of EMS appropriateness and protocol compliance evaluation performed by peer review. METHODS: Six peers retrospectively reviewed 168 patient care reports (PCRs) for severely injured trauma patients as defined by explicit criteria. Care was rated with a tool derived by a sequential derivation process. Emergency physicians (EPs) prospectively evaluated 118 of the patients with the same tool (blind to the PCR). Inter-rater reliability was determined between paramedics for all 168 PCRs. Criterion validity was determined between the EP (gold standard) and peer rating. Intra-rater reliability was determined from a repeated scoring of 50 PCRs after a washout period. The sample size was defined a priori. RESULTS: The criterion validity correlation coefficient was 0.29 (95%CI: 0.06-0.53). The inter-rater reliability kappa score was 0.35 (95%CI: 0.17-0.53). The intra-rater reliability kappa score was 0.62 (95%CI: 0.37-0.86). Three of 14 questions had very good inter-rater reliability: “Was a prehospital intervention required to manage airway?”; “Was a prehospital intervention required to cervical spine?”; and “Was a prehospital intervention required to orthopedic injuries?”. Two questions had strong criterion validity: “Was a prehospital intervention required to manage airway?” and “Was a prehospital intervention required to cervical spine?”. All questions asking if treatments provided were appropriate or compliant with written protocols had poor inter-rater reliability and criterion validity. CONCLUSION: Intra-rater reliability of retrospective peer review is good; however both inter-rater reliability and criterion validity are frequently poor. In particular, questions rating the appropriateness or compliance of treatments have poor reliability and validity. Although
more research is required to fully understand this issue, these results call into question the appropriateness of retrospective peer review of the quality of care in EMS.
Table of Contents

Abstract .................................................................................................................. ii

Table of Contents .................................................................................................. iv

List of Tables ........................................................................................................ vii

List of Figures ........................................................................................................ ix

Preface .................................................................................................................... xi

Preface ..................................................................................................................... xi

Thesis Objectives: ................................................................................................... xiii

Acknowledgements ................................................................................................. xv

1. Introduction ......................................................................................................... 1

2. Literature Review ............................................................................................... 5

   Epidemiology of Trauma ................................................................................. 5

   Quality of Care in Trauma ............................................................................. 5

   Protocols .......................................................................................................... 8

3. Methods ............................................................................................................. 9

   Stage 1: Development of Face Valid Data Collection Tool ......................... 9

      Derive Criteria ............................................................................................... 9

      Refinement of the Criteria ........................................................................ 11

      Weight Items ............................................................................................... 12

      Identify Items That Would Provide Little Information .............................. 12

      Select a Scoring System ............................................................................. 12

      Train Chart Reviewers .............................................................................. 12

   Stage 2: Data Collection .................................................................................... 13

      Case Enrolment ........................................................................................... 13
Preface

Since 1974, the British Columbia Ambulance Service (BCAS) has been the sole provider of prehospital Emergency Medical Services (EMS) in the province of British Columbia. BCAS serves a population of approximately 4.22 million people [1], covers an area of approximately 926,000 square kilometers, employs 3300 paramedics, and responds to approximately 500,000 calls a year [2]. It is one of the largest EMS systems in Canada.

Paramedics practice prehospital care under two different models of authority, which provide different forms of medical supervision. These two models are delegated practice and independent practice. In British Columbia, authority to practice as a paramedic is granted through licensure, providing paramedics an independent form of practice. Although, licensure provides a greater amount of autonomy for the practicing paramedic, this greater autonomy results in reduced medical supervision. Paramedics in BC are authorized to act without calling a physician and gaining authorization. Thus the paramedic must recognize what is wrong with the patient, and apply the appropriate treatment or protocol. The greater level of practice autonomy is viewed as a reasonable solution to the vast geography and potential communication problems that would exist if every paramedic, practicing in every community of BC had to contact a physician for authority to treat a patient [3]. However, this reduced level of medical supervision means that other methods of assessing and managing the activities of paramedics in the field must be found.

Quality improvement (QI) programs are one method to provide medical supervision and feedback to paramedics. BCAS does not currently have a formal QI
program that provides feedback to the paramedic. The closest form of practice feedback was the Re-certification Program, which was in place until 2002. Re-certification required that all Advanced Life Support (ALS) paramedics, and some Basic Life Support (BLS) paramedics (those who had attained the EMA II qualification) had to attend a two-week refresher course, followed by a written and practical exam, and continuing education. The remaining BLS paramedics (EMA I qualified) were required to maintain certification as Occupational First Aid Attendants, a program that is managed by the Workers’ Compensation Board of British Columbia. Since 2002, paramedics no longer are required to re-certify, but rather to attend annual Medical Education sessions. Systematic, organized evaluation of paramedic skill or practice no longer takes place.

Other QI activities within BCAS have focused on two areas: Emergency Medical Dispatch (EMD) and Key Performance Indicators. Three percent of all emergency calls to the BCAS communication centres are reviewed by a QI Coordinator. Six key areas of the calls are assessed and a cumulative score is assigned. [4] In the event of a major infraction of protocol, the QI coordinator discusses the call with the EMD in an attempt to improve EMD compliance.

The other area of QI for BCAS is the evaluation of Key Performance Indicators, which are measures of system performance such as financial management, performance to certain call time targets, staffing and training, and leave management. These data are measured for and reported to management, with no reporting out to paramedics.

Paramedics in BC are a large and essentially independent body of health practitioners; therefore, it is surprising that they have received so little focus for QI activities, especially in the absence of direct or indirect medical supervision. There are
several reasons for having an interest in evaluating the quality of paramedic practice: increasing understanding of the impact of prehospital care on patient outcomes, understanding the different needs for education and training in the varied geographic and demographic settings of the provinces, and evaluating patient safety within the EMS system. Evaluating the quality of care provided by paramedics requires valid and reliable measures of both patient outcomes and care processes.

**Thesis Objectives:**

There were three objectives for this thesis:

1. Develop a content valid data abstracting form to be used to evaluate protocol compliance and treatment appropriateness for trauma patients;
2. Quantify the intra-rater, and inter-rater reliability and the criterion validity of the data abstracting form;
3. When it was discovered that the data abstracting form was neither reliable nor valid, it was decided to explore and quantify the potential sources of error that may have contributed to poor reliability and validity.

This thesis is structured into 6 chapters using the following framework. Chapter One describes a conceptual framework by which quality of care is commonly evaluated. Chapter Two is a review of the literature related to EMS care of trauma patients. Chapter Three describes the methods used to create the data collection tool. Chapter Four reports the findings of the validation process, specifically reporting the reliability and validity of the form, and an exploration of the potential sources of error. Chapter Five, discusses the findings within the context of EMS, and where literature is lacking specifically on EMS systems, refers to other areas of medicine.
The preference was to remain within EMS first, and trauma second, but where this was not possible, other areas of medicine were consulted. Chapter Six discusses the policy implications and gives suggestions for further research.
Acknowledgements

It would be an understatement to say that many people have provided me assistance with this thesis. In fact, everyone I have worked with or encountered throughout my time in the Department of Health Care and Epidemiology at the University of British Columbia has affected me profoundly. But there are a few people who warrant special acknowledgement. First, I would like to thank all of the physicians who provided their limited time and expansive expertise during the initial developmental phase of the data abstraction form: Drs Iain MacPhail, and Joseph Ip, of the Royal Columbian Hospital, Drs David Harrison, and Richard Simons of Vancouver General Hospital, and Drs Robert Stenstrom and Devin Harris of St Paul’s Hospital in Vancouver. Next I would like to thank the physicians and paramedics who lent me their time and thoughts in the interview phase of this project: Drs Karen Wanger and Jim Christenson of St Paul’s Hospital, and Ian Macmillan, and Andy Fletcher of BCAS in Vancouver. Every research project has several people working in the background doing very important administrative tasks and keeping things in order: Catherina Van Beek, Susan Goegen, and Jan Buchanan were all extraordinary research nurses. I would also like to thank the 6 raters who provided me with the data, and all of the BC Ambulance Service Paramedics who listened to my ideas, and challenged me to keep my project relevant.

My thesis committee has been especially wonderful. Each of them brought a different and very important perspective to this project: Dr Bruno Zumbo as the expert psychometrician. It has been an honour to be counted as one of Dr Zumbo’s students. Dr Robert Reid whose course and ongoing challenges have been greatly appreciated. Dr Adrian Levy who started my journey into measurement, and continue to challenge me to
make my work explicit and clear. Dr Riyad Abu-Laban who provided me inspiration and guidance for much longer than this project, and I am very happy that you agreed to become a formal part of the committee. And finally, Dr Sam Sheps: whenever anyone asks me whom I would recommend as a supervisor, your name comes. References from Grad students might just be the greatest honour one can bestow upon a supervisor. You have been an invaluable mentor and teacher. And I hope you will accept me back when that time comes.

The greatest acknowledgement goes to my wife, Barb. She unwittingly agreed to this project, which has kept me away from her for nearly 4 years. Thanks for your sacrifice. I love you and appreciate your patience and understanding.
1. Introduction

Background and Significance

Donabedian conceived a framework for assessing quality of care [5] which required the evaluation of system structures, processes, and outcomes. Structure refers to the structural components of a medical system, such as equipment, training, and standards or certification.

Examples of structural components in Emergency Medical Services (EMS) are dispatch systems, paramedic qualifications, and continuing education programs. A dispatch system might enhance quality through dedicated and secure telephone lines, and an easily remembered phone number such as 911, and a set of instructions for conducting Cardio-Pulmonary Resuscitation (CPR) over the telephone. Improved patient outcomes would result from being able to identify the patient’s location rapidly and accurately, passing that information on to the closest ambulance, and providing any necessary instructions to bystanders by phone.

Process refers to what care-givers do in providing care to patients within the system. An example of a process component is “response time” which is how quickly an EMS agency responds to a call for assistance. Improved patient outcomes result from reducing the time between the occurrence of an injury and the arrival at a hospital for definitive treatment.

Outcomes refer to a measure of the patient’s status as a result of his encounter with the medical system. For example, a patient who suffers a pulmonary contusion from a motor vehicle collision (MVC) has some chance of dying as a result of this injury. The purpose of the EMS system is to locate the patient, get to them rapidly, treat the injuries,
and transport them to the hospital. An MVC patient’s status at the end of the EMS system encounter, usually hospital discharge or death, is considered an outcome. Table 1 displays many other structure, process and outcome measures that can be used to evaluate the quality of EMS.

Table 1.1 Examples of Structure, Process and Outcomes in EMS

<table>
<thead>
<tr>
<th>Structure</th>
<th>Process</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispatch system</td>
<td>Accuracy</td>
<td>Survival</td>
</tr>
<tr>
<td>Deployment</td>
<td>Time intervals (&lt; 8 min)</td>
<td>Length of Stay</td>
</tr>
<tr>
<td>Equipment</td>
<td>Compliance</td>
<td>ICU Length of Stay</td>
</tr>
<tr>
<td>Protocols</td>
<td>Success rates</td>
<td>Complications</td>
</tr>
<tr>
<td>Training/Qualification</td>
<td></td>
<td>Cost</td>
</tr>
<tr>
<td>Trauma Registry</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Process, or what happens at the point of patient contact, is a valuable source of information when evaluating quality of care. Process in itself can be a study outcome, if the relationship between the structural component and patient outcome is well established. For example, defibrillation of a lethal cardiac arrhythmia ventricular fibrillation is known to trigger return spontaneous circulation in some patients. Therefore, accurate diagnosis of ventricular fibrillation by paramedics may be selected as a valuable intermediate outcome in an EMS system review of its quality. If paramedics are unable to accurately diagnose ventricular fibrillation, then they will likely not know when to appropriately defibrillate patients who present with this condition.

Donabedian’s structure, process and outcomes model provides a framework for theorizing and testing the causal relationship between each of these components. For example, a treatment protocol (structure) that is complied with (process) by practitioners
should improve survival or reduce lengths of stay in the hospital (outcomes). Therefore, EMS systems have equipped responders with automatic external defibrillators (AED), trained responders to use them, and observed improved cardiac arrest survival. To test quality at the system level, medical supervision staff will then be interested in tracing the availability of the defibrillator (structure) on the ambulance, through to the accurate recognition of patients in ventricular fibrillation (process) and the return of spontaneous circulation (outcome) in their evaluation of the quality of their EMS system.

Sources of Data Used to Measure Quality

Quality is commonly measured retrospectively by having peers review medical records. Despite being a widespread method for measuring quality in EMS, retrospective peer review is not well studied. It is common for EMS systems to use paramedics in as Quality Improvement (QI) officers. QI officers spend much of their time doing data abstraction from patient care reports (PCR). The abstracted data are used to evaluate the quality of the system. Few EMS system personnel have a good understanding of the quality of their data. Commonly, data quality refers to the completeness of the medical record. Quality rarely refers to whether the data represent the veracity of what is being measured, or if there are differences in the way the data are recorded based on how it is collected or by whom it is collected.

Accurate measurement of the structure, process and outcomes of the EMS system could reduce misclassification bias of both exposure (structure and process) and outcomes. It is important to ensure that valid and reliable measurement techniques are used in research, and QI activities. Invalid and unreliable measurement techniques can result in misdiagnosis of a problem and deployment of resources directed at the wrong
problem. Therefore, if quality improvement programs are intended to benefit patients, valid and reliable measures of the system must be developed.
2. Literature Review

Epidemiology of Trauma

Trauma due to unintentional injury is the sixth-leading cause of death of British Columbians of all ages and the leading cause of death of British Columbians between the ages of one and forty-four. Males account for two-thirds of the 1621 lives lost due to trauma in 2003 [6]. Approximately 424,000 people are injured each year, with 27,000 hospitalized, and 9000 partially disabled [7]. The estimated economic burden of unintentional injuries is $2.1 billion dollars annually. Trauma resulted in the second highest number of potential years of life lost [7]. Approximately 32% of the calls BCAS responded to were traumatic injuries, meaning that approximately 160,000 of the individuals attended to by BCAS had suffered some form of trauma [2].

The trauma mortality profile in British Columbia is similar to the U.S. profile [8], and strategies to reduce morbidity and mortality associated with trauma in British Columbia are modeled after U.S. programs. The British Columbia Ambulance Service (BCAS) treats injury victims according to the American College of Surgeon’s (ACS) Committee on Trauma (COT) recommendations [9] [10]. These include a set of triage rules and treatment protocols that are intended to reduce morbidity and mortality in some trauma patients.

Quality of Care in Trauma

The time between a 911 call and the patient’s arrival at hospital is by far the most commonly measured process variable in studies that evaluate the quality of care in EMS. Several studies [11-15] measured time as the outcome of interest implying or directly stating that time is an indicator of EMS system performance. The measurement of time as
an outcome without considering its effect on patient mortality or morbidity implicitly reinforces the “the scoop and run” philosophy of prehospital care without providing any direct evidence to support the position that shorter prehospital times do indeed reduce morbidity or mortality. In an attempt to better understand the relationship between time and trauma patient outcomes, Sampalis and colleagues [16] used time as an intervening variable in a multivariate analysis. Their finding was that prehospital time, which was highly correlated with the number of prehospital treatments, contributed to increases in mortality. Lerner found, on the other hand, that mortality was not consistently correlated with prehospital time, and speculated that paramedics recognize and triage the most severely injured patients, those with the highest Injury Severity Scores (ISS), faster than less injured patients [17]. Although the debate over “stay and play” versus “scoop and run” has long been debated in the medical literature, evidence of a beneficial or detrimental effect of increased prehospital times does not appear to be firmly established in the scientific literature.

Another commonly measured quality of care process is completeness of documentation. A number of studies sought to describe the quality or quantity of documentation as a direct measure of the quality of care [13, 18-20]. Only one study sought to establish a relationship between the completeness of prehospital documentation and some other variable[21]. The relationship established by Burstein and colleagues was that incomplete documentation results in a biased estimate of the probability of mortality based on the mechanism criteria for trauma triage guidelines. No studies attempted to demonstrate that poor documentation affected patient outcomes although several theoretical relationships could be tested including that between poor
documentation and increased medical errors, delays in treatments, or reduced continuity of care from the prehospital segment to the in hospital segment of care.

Appropriateness of care or compliance with protocols or guidelines has rarely been studied although it is commonly advocated [22]. Several studies evaluated either appropriateness of care, or compliance with protocols as the primary study outcome [23-27]. By far the most commonly evaluated protocol is the trauma diversion protocol. Trauma diversion is a directive to take patients who meet certain explicit criteria directly to a trauma-receiving centre, in some instances bypassing a closer hospital. Ma and colleagues evaluated compliance with triage guidelines to evaluate which system or patient characteristics were responsible for non-compliant treatments. The support for trauma diversion protocols has been established through before-and-after studies of the introduction of diversion programs or the introduction of regional trauma programs [23]. Few other explicit process of care measures have been evaluated for prehospital trauma, yielding little useful data for quality improvement programs other than that an inappropriate treatment process was performed.

Helm and colleagues [28] evaluated the quality of prehospital respiratory support by evaluating blood gases of brain-injured patients who had been intubated in the prehospital setting. This study is one of the few studies of prehospital trauma care that attempted to establish some connection between care provided by paramedics and a clinical outcome. Their findings were that paramedics frequently failed to provide optimal or acceptable oxygenation of brain injured patients. Although this study should be applauded for evaluating the effect of EMS care process, it failed to draw the connection between poor patient care (i.e. poor prehospital oxygenation for severe brain
injuries), and a relevant patient outcome such as death or level of neurological function at hospital discharge.

EMS is well established in North America. However, it is poorly studied [29]. Because it is well established, experimental designs may not be an option for evaluation of system performance. Observational studies are susceptible to bias and confounding. This means that the ability to draw connections, specifically cause and effect relationships between EMS system structures (such as ALS interventions for severely injured patients) and patient outcomes (such as neurological scores at hospital discharge) is confounded by the process of care (specifically the paramedic’s adherence to protocol). Therefore, valid and reliable measurement of care processes is warranted.

Protocols

Prehospital trauma protocols are physician orders directing paramedics to treat patients in a prescribed manner[30-31]. With this in mind, both physicians and BCAS QI personnel should carefully scrutinize the care delivered in the prehospital setting to ensure that prescribed treatments are followed. Unfortunately, validated criteria and data capture instruments are lacking, which can result in an approach to QI activities, which is time and resource consuming, and arbitrary. Lack of standardized, and explicit data collection techniques can result in misclassification bias, and spurious inferences. Data collection using a standard and explicit tool could provide valuable information about the process of prehospital trauma care which would assist providers, planners, Emergency Physicians and Trauma Surgeons in improving the quality of trauma care in the Province of BC.
3. Methods

This project was designed to progress through three main stages:

1. Development of a face valid data collection tool

2. Data collection
   a. Enrolment of cases
   b. Collection of ratings using the data collection tool

3. Quantification of Reliability and Validity
   a. Calculation of agreement and correlation statistics
   b. Conceptual and empirical exploration of potential sources of error.

Stage 1: Development of Face Valid Data Collection Tool

Ashton and colleagues [32] described the Criteria Development Method to develop explicit process of care criteria. With slight modifications, 5 of the 6 steps outlined by Ashton and colleagues have been used here.

Ashton’s recommended steps are the following:

∞ Derive criteria with which to evaluate process of care

∞ Refine criteria, and create items with which to perform evaluation

∞ Weight items

∞ Identify items that will provide little information

∞ Selection of scoring method

∞ Train chart reviewers.

Derive Criteria

Ashton suggests collecting the “state of the art literature” to ensure that the criteria being used are up to date. For this study, there was no flexibility for practitioners
to practice outside of protocols established by BCAS, so the materials collected were the current training literature, and BCAS Protocols. These materials were distributed to a panel of 6 emergency physicians and 1 trauma surgeon, and are included as Appendix 1. Care was taken to ensure the material was complete and current. Based on these materials and clinical experience, a list of potential measurement domains (specific things to be evaluated), and several questions for each domain were created.

The seven panel members who each work in one of the three trauma centres in the Greater Vancouver Regional District assisted with the development of the data abstraction form. Table 3.1 displays the clinical qualification and duration of practice for each of the panelists. Each panelist was to preview all of the BCAS protocol. The objective of this focus group was to create an exhaustive list of items and generate appropriate responses. The expert panel met in person for one two-hour meeting at which they discussed the proposed domains, reviewed the proposed questions, added questions, and provided responses to the questions.

**Table 3.1 Expert Panel Qualification, Practice Duration and Hospital**

<table>
<thead>
<tr>
<th>Physician</th>
<th>Clinical Qualification</th>
<th>Duration of Practice</th>
<th>Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ABEM, FRCP(C)</td>
<td>7 years</td>
<td>Vancouver General Hospital</td>
</tr>
<tr>
<td>2</td>
<td>CCFP(EM), FRCP(C)</td>
<td>13 years</td>
<td>Vancouver General Hospital</td>
</tr>
<tr>
<td>3</td>
<td>CCFP(EM)</td>
<td>2 years</td>
<td>St Paul’s Hospital</td>
</tr>
<tr>
<td>4</td>
<td>FRCP(C)</td>
<td>1 year</td>
<td>Royal Columbian Hospital</td>
</tr>
<tr>
<td>5</td>
<td>FRCS, FRCS(C), FACS</td>
<td>12 years</td>
<td>Vancouver General Hospital</td>
</tr>
<tr>
<td>6</td>
<td>CCFP(EM)</td>
<td>13 years</td>
<td>Royal Columbian Hospital</td>
</tr>
<tr>
<td>7</td>
<td>CCFP(EM)</td>
<td>3 years</td>
<td>St Paul’s Hospital</td>
</tr>
</tbody>
</table>
Refrinement of the Criteria

The questions and responses were organized into the domains of prehospital trauma care and this list was emailed to the panelists. Wishing to ensure full participation from the group, Ashton opted for a Delphi Process to refine the criteria. A Delphi Process [33] was deemed unnecessary for two reasons: criteria or standards for treatment had already been established and documented in the training material and BCAS policy, and the panelists were not certain of the criteria to be used; an alternative, the Content Validity Ratio [34] was available and would provide much the same results while imposing minimal time demands on the panelists.

Refrinement was performed through email iterations, where the expert panelists rated the items developed as “essential”, “useful, but not essential”, or “not necessary” following the methods described by Lawshe [34]. The content validity ratio (CVR) is a ratio of the number of essential items minus half of all items over half of all items. It is derived as follows:

\[
\text{CVR} = \frac{n_e - N/2}{N/2}
\]

Where, \( n_e \) is the number of essential items, and \( N \) is the total number of items.

Items were retained if the value of the CVR exceeded the level of chance concurrence at the 5% level. In the case of this panel of seven experts, minimum value of the CVR was 0.78 [34]. From this process, two data collection forms were generated: the first for the prospective collection of data performed by emergency room physicians, and the second for the retrospective review to be performed by the peer reviewers.
**Weight Items**

Item weights were not derived. In the event that a valid and reliable instrument had been developed, empirical weights would have been provided.

**Identify Items That Would Provide Little Information**

Identifying items that would provide little information was primarily achieved through the calculation of the CVR. Concurrent to item reduction based on the CVR, the four interviews were conducted with two experienced EPs and two experienced Advanced Life Support (ALS) paramedics to get a diversity of input and perspectives on the proposed items. The interviewee provided their interpretation of the question and the responses and identified any ambiguous wording. The notes from these interviews were reviewed and any consistent themes or ambiguous wording was used to clarify the question.

**Select a Scoring System**

The expert panel decided that two response types would be used. For questions that lacked an explicit protocol, the item responses were “optimal”, “acceptable”, and “not appropriate”. For questions with an explicit protocol, the item response was “compliant” and “not compliant”.

**Train Chart Reviewers**

Instructions on completing the data abstracting form were provided in writing to each of the peer raters. Each peer was instructed to complete the review within 48 hours of starting. This was done to improve the internal consistency of the ratings. All educational and protocol material was provided as reference materials. (Refer to Appendix 1.)
Stage 2: Data Collection

Case Enrolment

There were two sources of cases for this study. First, patients were enrolled prospectively through the Emergency Department (ED) at Vancouver General Hospital (VGH). Starting in December 2003, all trauma patients who met the following criteria were to be reviewed by the attending EP:

1. Scene response by BCAS with direct transport to VGH (i.e. no secondary transfers). This condition ensured that the EP observed the prehospital care, and that the original BCAS crew report could be located for retrospective review by peer raters; and

2. Met BCAS Trauma 99 definition for major trauma; and

3. Were referred for consult to any of the following hospital services: Trauma, General, Orthopedic, or Neurological Surgery Consult, or the patient remained in the ED for greater than 6 hours; and

4. Cases were included when all of the previous conditions were met, and the attending EP completed a data collection form within 48 hours of examining the patient.

The second source of cases was the British Columbia Trauma Registry (Trauma Registry). Preliminary calculations of patient enrollment through VGH ED estimated it would take up to 8 months to collect the minimum number of patients required for this study. The decision was made to sample from the Trauma Registry as an additional source of cases, which would be used to evaluate intra-rater reliability. Therefore a random sample of 60 patients who were included in the Trauma Registry and had a
hospital discharge date between 01 April 2002, and 31 March 2003 and who met the following characteristics were included:

1. Injury Severity Score ≥ 16;

2. Between the ages of 16 and 64; and

3. Transported to hospital by BCAS.

These criteria for inclusion were selected to ensure that only adults who had suffered major trauma were included. Although the treatments provided to children and less severely injured patients are the same as those performed on adults, the primary interest of this review was to discover determinants, which would reduce morbidity and mortality in adults. Less severely injured patients, those with Injury Severity Scale scores of less than 15, generally suffer less death and disability than those with higher scores [36].

The final criterion for eligibility as a case for both the prospective enrollment and retrospective enrollment was the availability of the original BCAS crew report. If the crew report was not available, then the case was excluded from the study. A total of 168 cases were enrolled: 118 from the emergency department and 50 from the Trauma Registry. Figure 3.1 shows the two case enrolment categories and how they relate to the type of measurement considered.

Collection of Ratings

There were three ratings, which provided data for the measurement of validity and reliability. First, an EP evaluated each patient arriving at the VGH ED who met the

---

1 ISS >15 is a generally accepted cut-off for major trauma patient quality improvement activities 35. Bruns B, Trauma System Quality Assurance, in Quality Management in Prehospital Care, Swor RA, Editor. 1993, Mosby Lifeline: St Louis, MO. p. 161 - 187. Use of a lower cut-off would result in capture of patients with less severe injuries that are not the target of most trauma systems.
prospective rating enrolment criteria described above. The EP performed a physical examination of the patient and listened to the paramedic's verbal report. Based on the information obtained the EP completed the physician version data abstracting form. The physicians were instructed not to review the BCAS patient care record (PCR) prior to completing the data abstracting form, to ensure that their only source of information was their physical exam and the paramedic's verbal report.

All cases (those enrolled in the ED, and those enrolled from the Trauma Registry) were reviewed by 6 peer reviewers. After a 3-month washout period, a sub-set of 50 cases, those enrolled from the trauma registry, were reviewed and rated a second time by the peers. Based on these ratings, criterion validity, inter-rater and intra-rater reliability were evaluated. Figure 3.1 shows the details of case enrolment and its relationship to the type of measurement performed.
Figure 3.1 organizes the completion of the data abstracting form into three time periods, displayed in the columns. This figure also displays the type of measure being performed based on the rows or columns. From December 2003 to February 2004, EPs prospectively enrolled and rated patients arriving at the ED in VGH. From July to September 2004, peer raters completed data abstracting forms for all cases enrolled from the ED and the Trauma Registry. To reduce recall which results in a spurious inflation of intra-rater reliability, a 3-month washout period was provided for the peer reviewers. From October to December 2004, paramedics reviewed the 50 cases from the Trauma Registry.
Step 3: Quantification of Reliability and Validity

Intra-rater Reliability:

Unweighted kappa scores [37] were calculated to measure agreement between the first and second rating of the Trauma Registry cases. This is displayed as the second row in Figure 3.1. In the event that an unweighted kappa score could not be calculated because of a non-square contingency table\(^\text{a}\), a non-square kappa statistic was calculated based on a macro provided by SPSS Inc [38].

Two sets of mean kappa statistics were generated to measure inter-rater reliability. The original intention of this thesis was to measure agreement on the assessment of trauma care. Upon review of the data abstracting tools, it became apparent that the raters were not always able to assess the care provided, and left many of the questions blank or responded with “cannot tell”, even when this was not an option. Therefore, the first set of kappa statistics presented are those calculated with the responses “cannot tell” and “N/A” excluded. It is proposed that these kappa statistics measure agreement on the quality (i.e. compliance with protocols, or appropriateness of treatments). A second set of mean kappa statistics were calculated with “cannot tell” and “N/A” included. These responses were intended to evaluate both agreement on the quality of the documentation and care provided. Missing values or blank responses were not included in any of the analyses. These responses appeared on the data collection forms where the rater had included the option of “cannot tell”, so there was no way of knowing precisely what was intended with this response.

\(^\text{a}\) Non-square contingency tables result when one or more cells are empty or equal to zero.
**Inter-rater Reliability**

All 168 cases were included in the evaluation of agreement among the raters. This is displayed as the middle column of Figure 3.1. To calculate unweighted kappa statistics for all raters a multiple rater kappa macro, available from SPSS Inc.[38] was used. This macro is based on the work of Seigel and Castellan[39].

**Kappa Statistic considerations**

Kappa statistic is a chance-corrected measure of agreement originally derived by Cohen [40]. Kappa is the most commonly reported measure of agreement, and is considered superior to reporting raw agreement because the statistic considers and corrects for agreement expected by chance[41]. Only unweighted kappa statistics have been calculated and reported here.

Cohen’s kappa calculates agreement between two observers. In this study, we are interested in the agreement among 6 observers and over 14 questions. Siegel and Castellan point out that kappa statistics are approximately normally distributed with a mean of 0 [39]; therefore significance testing was based on the normal distribution. Kappa derived using Siegel and Castellan’s method is the primary statistic reported for inter-rater reliability. Test for differences between raters and questions has been carried out using similar assumptions. Ninety-five percent confidence intervals (CI 95%) were calculated for all mean kappa statistics.

Landis and Koch [42] divide kappa statistics into six categories, specifying that a kappa statistic of 0.61 to 0.80 represents substantial agreement, and 0.81 to 1.00 represents near perfect agreement. Fleiss [37] suggests that kappa statistics greater than
0.75 represent excellent agreement beyond chance. For this thesis, 0.70 was the cut-off for inclusion in the final instrument.

**Criterion Validity**

With emergency physician ratings as the criterion measure, correlation between the peers and the physicians was calculated. For questions involving a binary response (yes and no, or compliant and not compliant) a phi coefficient was calculated. For ordinal questions involving an evaluation of the appropriateness of the treatments, a Spearman’s correlation was calculated. Cut-off for inclusion in the final version was 0.70.

**Conceptual and Empirical Investigation of Potential Sources of Error**

Upon review of the kappa statistics and correlation coefficients, it was discovered that only two questions had surpassed a 0.70 cut-off. To test for the effect of the EP prospective ratings on the quantity of the correlations, one final post-hoc episode of data collection was performed. The purpose of collecting these data was to evaluate if a lack of a homogeneous gold standard ratings had had an effect on the validity correlations.

The principal investigator and one of the members of the supervising committee (RBA) rated a random sample of 19 forms. The ratings were derived by consensus, wherein the two raters reviewed each question together, discussed the question, and came to consensus on the rating. This validation subset was then used as the criterion measure for the calculation of a new set of phi or Spearman’s Correlation coefficients.

**Phi and Spearman Correlation coefficients**

Zimmerman and colleagues [43] expanded on the work from the early 20\textsuperscript{th} century by R.A. Fisher regarding the distribution characteristics of correlation coefficients.
Zimmerman and colleagues provided formulas to transform the correlation coefficient into a Z score, performed the back transformation from the Z score into the correlation coefficient, and derived a standard deviation for a Z score. To perform comparisons between raters and questions, all correlation coefficients were first transformed into Z scores. A mean Z score was calculated for each rater (all questions) and for each question (all raters). A standard deviation was then calculated and confidence intervals derived for each mean score by question and by rater. Finally the Z to r transformation was performed for the lower and upper confidence interval and for the mean Z score producing unbiased mean correlation coefficients and confidence intervals.

**Statistical Considerations and Sample Size Calculation**

Two aspects of this study require consideration for the sample size calculation. These are the number of cases enrolled and the number of repeat observations. Traditionally, precision of parameter estimates is the prime concern when considering sample size. However, budget constraints has been established as a valid criteria provided meaningful results can be anticipated with the sample size proposed [44 - 45]. For this thesis, a time frame budget of 8 months was established *a priori*, and was then used to determine the precision of the confidence intervals derived from the study data. The following are the details of the budget and sample size considerations. Based on a VGH Trauma Services estimate of 18 trauma team activations per month for the VGH ED, and an estimate (based on experience with similar projects) that 75% of these activations would be rated by EPs, an eight-month period of data collection would result in 108 trauma events. Assuming good correlation (correlation value of 0.70) between physician and paramedic ratings, this number of events would result in confidence
intervals of ± 0.10, a degree of precision deemed reasonable by the study team[46]. Using the simplified formula for the determination of the number of repeated ratings under a scenario of budget constraints, [44] 18 separate ratings would be performed on the 108 trauma cases for a total of 1944 cases rated. This resulted in three ratings, one by the EPs and two by the peers. This was operationalized as one prospective rating by the EPs and 2 retrospective ratings by six peers.

Ethics, Privacy and Confidentiality

The UBC Clinical Review Ethics Board (CREB) required that the identities of all study subjects be protected. Therefore, the identities of the trauma patients who were enrolled in the study, the paramedics who provided the treatments to the patients, and the physicians and peers who observed and evaluated the care needed to be guarded and kept private. To accomplish this each BCAS patient care report was masked. This included masking the names, date of birth, personal identifying numbers (such as Provincial Health Number, and employee or paramedic licensing number), signatures, addresses (both billing address and BCAS response address), Response or Event numbers, and event dates. All of the billing fields, crew identifier fields, and the body of the crew reports were searched. Once the original crew reports were masked, copies were made so that unmasking (removal of paper or tape) could not occur.

To reduce the possibility of a Hawthorne Effect, a “deceive and debrief” strategy was employed. It was suspected that if paramedics knew that their documentation was being reviewed for protocol compliance and appropriateness of treatments, they would take greater care with documentation of ambulance call details. This change in behavior
was to be avoided. Therefore, paramedics (outside of those involved as raters) were not informed of the study.

This study received ethical approval from the Clinical Ethic Review Board of UBC. This study also received approval for the use of a "deceive and debrief" strategy from the Ambulance Paramedics of British Columbia (CUPE 873) the collective bargaining unit for paramedics in British Columbia.
4. Results

Development of Face Valid Data Collection Form

*Measurement Dimensions*

Six dimensions of EMS trauma care were identified after review of BCAS Policy and Procedure. These were airway maintenance, breathing maintenance, circulation support, cervical spine management, orthopedic injury management, and trauma diversion[47]. These dimensions represent the treatment priorities for paramedics when they encounter an injured person. Airway maintenance refers to the need to establish and maintain the patency of a person’s airway by any of the following: manual manipulation, suctioning, endotracheal intubation, or placement of an airway adjunct. Breathing maintenance refers to the use of oxygen delivery equipment or manual ventilation of the patient in respiratory arrest or distress. Circulation support refers to either the performance of cardio-pulmonary resuscitation when a person is in cardiac arrest, or the provision of intravenous fluids when the patient has a low blood pressure. Cervical spine management refers to the use of manual stabilization when a patient injures their head, neck, or back. Extremity injuries also often occur, requiring manual stabilization or pain relief measures to be taken. Finally, patients who meet specific criteria are required to be transported to a trauma receiving hospital, which entails diverting away from a non-trauma hospital.

Of these dimensions, explicit protocols were available for three: circulation support, pain management (orthopedic injury management), and diversion. Implicit guidelines were available for the other four dimensions, although these were somewhat
superficial in nature. For example, a general guideline exists that states airway must be managed prior to breathing in a trauma patient.

**Expert Panel Activities**

Two EPs from each of the trauma receiving facilities in the lower mainland, and a trauma surgeon from VGH were included in the expert panel, which met to add questions, provide responses, and clarify the areas to be measured. Twenty-four questions were suggested in this session, with item responses.

**Item Reduction**

Following the expert panel meeting, each member of the expert panel received two copies of the questions: one that would become the form used by EPs as patients were enrolled at VGH, and one that would be used by paramedic raters. Each expert panel member rated each question as “essential”, “useful, but not essential”, or “not necessary”. Based on the formula provided by Lawshe, [34] the Content Validity Ratios (CVR) were calculated for each of the questions. A minimum score of .99 was required for a question to remain in the data abstracting form.

**Interviews**

Two EPs, not involved in the study, and two paramedics were interviewed concurrent to the expert panel item reduction exercise. The interviewees were presented with the data abstraction form, and asked to comment on the proposed questions. They were invited to clarify language or suggest changes or additions. Only suggestions that were provided by all interviewees were included in the final edition of the form. However, no changes were made to the form as a result of these interviews. Detailed notes were made, and these are included in Appendix 2.
Physician Instrument

There were 24 questions proposed during the expert panel meeting. Only 13 of these were included in the final version of the data abstracting form based on the email responses from the panel. Table 4.1 displays the proposed questions, and the CVR for each.
<table>
<thead>
<tr>
<th>Physician Instrument</th>
<th>Question</th>
<th>CVR Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Airway</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there a need to manage this patient’s airway in the prehospital setting?</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Indicate reason for need to manage airway: Respiratory Failure; Failure of gas exchange; Airway Patency; Airway Protection; Other indications</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>How was this patient’s airway managed? Manually; Oral Airway; Endotracheal Intubation; Cricothyrostomy; Other</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>The paramedic’s management of this patient’s airway was: Optimal; Acceptable; Not Appropriate; Cannot tell</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Respiration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there a need to manage this patient’s breathing in the prehospital setting?</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Indicate reason for need to manage breathing: Respiratory distress; Oxygenation; Other Indication</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>How was this patient’s breathing managed? Oxygen ____ lpm; BVM without intubation; BVM with intubation; Needle decompression; Other</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>The paramedic’s management of this patient’s breathing was: Optimal; Acceptable; Not Appropriate; Cannot tell</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Circulation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there a need to manage this patient’s blood pressure in the prehospital setting?</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Indicate reason for need to manage this patient’s blood pressure: Hypovolemia; Systemic BP &lt; 90; Signs of shock; Burns to &gt; 20% BSA; Other Indication</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>How was this patient’s blood pressure managed? IV Solutions; Number of IV Cannula; Amount of fluid; Crew not IV endorsed</td>
<td>0.43</td>
<td></td>
</tr>
</tbody>
</table>
The paramedic's management of this patient's blood pressure was: Compliant with BCAS Protocol; Not compliant with BCAS Protocol

<table>
<thead>
<tr>
<th>Cervical Spine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there a need to manage this patient's cervical spine in the prehospital setting?</td>
</tr>
<tr>
<td>Indicate reason for need to manage patient's cervical spine: LOC; Mechanism of Injury; Injury to head/neck; Associated Injuries; Other indication</td>
</tr>
<tr>
<td>How was this patient's cervical spine managed? Cervical Collar; Backboard; Clamshell; Other</td>
</tr>
</tbody>
</table>

The paramedic's management of this patient's cervical spine was: Optimal; Acceptable; Not appropriate; Cannot tell

<table>
<thead>
<tr>
<th>Other Orthopedic Injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence of injuries to patient's extremities? Yes; No</td>
</tr>
<tr>
<td>Was there a need to manage this patient's extremity injuries in the prehospital setting?</td>
</tr>
<tr>
<td>How was this patient's extremity injuries managed? Splint; traction splint</td>
</tr>
</tbody>
</table>

The paramedic's immobilization of this patient's extremity injuries was: Optimal; Acceptable; Not Appropriate; Cannot tell

<table>
<thead>
<tr>
<th>Code 99</th>
</tr>
</thead>
<tbody>
<tr>
<td>The patient required diversion to a trauma facility based on: MOI; Anatomical findings; Physiologic presentation</td>
</tr>
</tbody>
</table>

The paramedic Comply; Did not Comply; with Code 99 protocol

A copy of the physician form follows as Figure 4.1
### Figure 4.1 Final Version of Physician Data Abstracting Form

<table>
<thead>
<tr>
<th>Airway</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was there a need to intervene to manage this patient’s airway in the prehospital setting?</td>
<td>1 _ Yes</td>
<td>2 _ No</td>
</tr>
<tr>
<td>2. The paramedic’s management of this patient’s airway was:</td>
<td>1 _ Optimal</td>
<td>4 _ Cannot tell</td>
</tr>
<tr>
<td></td>
<td>2 _ Acceptable</td>
<td>5 _ N/A</td>
</tr>
<tr>
<td>3 _ Not Appropriate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Breathing</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Was there a need to intervene to manage this patient’s breathing in the prehospital setting?</td>
<td>1 _ Yes</td>
<td>2 _ No</td>
</tr>
<tr>
<td>4. The paramedic’s management of this patient’s breathing was:</td>
<td>1 _ Optimal</td>
<td>4 _ Cannot tell</td>
</tr>
<tr>
<td></td>
<td>2 _ Acceptable</td>
<td>5 _ N/A</td>
</tr>
<tr>
<td>3 _ Not Appropriate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Circulation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Was there a need to intervene to manage this patient’s blood pressure in the prehospital setting?</td>
<td>1 _ Yes</td>
<td>2 _ No</td>
</tr>
<tr>
<td>6. The paramedic’s management of this patient’s blood pressure was:</td>
<td>1 _ Compliant with BCAS Hypovolemia Protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 _ Not Compliant with BCAS Hypovolemia Protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 _ N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cervical Spine</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Was there a need to intervene to manage this patient’s cervical spine in the prehospital setting?</td>
<td>1 _ Yes</td>
<td>2 _ No</td>
</tr>
<tr>
<td>8. The paramedic’s management of this patient’s cervical spine was:</td>
<td>1 _ Optimal</td>
<td>4 _ Cannot tell</td>
</tr>
<tr>
<td></td>
<td>2 _ Acceptable</td>
<td>5 _ N/A</td>
</tr>
<tr>
<td></td>
<td>3 _ Not Appropriate</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extremity Injuries</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Was there a need to intervene to manage this patient’s extremity injuries in the prehospital setting?</td>
<td>1 _ Yes</td>
<td>2 _ No</td>
</tr>
<tr>
<td>10. The paramedic’s immobilization of this patient’s extremity injuries was:</td>
<td>1 _ Optimal</td>
<td>4 _ Cannot tell</td>
</tr>
<tr>
<td></td>
<td>2 _ Acceptable</td>
<td>5 _ N/A</td>
</tr>
<tr>
<td></td>
<td>3 _ Not Appropriate</td>
<td></td>
</tr>
<tr>
<td>11. The paramedic’s management of this patient’s extremity injuries was:</td>
<td>1 _ Compliant with BCAS Pain Management Protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 _ Not Compliant with BCAS Pain Management Protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 _ N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code 99</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The patient required diversion to a trauma facility based on:</td>
<td>1 _ Mechanism of injury</td>
<td>2 _ Physiologic Presentation</td>
</tr>
<tr>
<td>13. The paramedic was:</td>
<td>1 _ Compliant with Code 99 Protocol</td>
<td>3 _ N/A</td>
</tr>
<tr>
<td></td>
<td>2 _ Not Compliant with Code 99 Protocol</td>
<td></td>
</tr>
</tbody>
</table>
Guide to Questions

### Airway: Questions 1 & 2
The BCAS protocols for airway intervention are minimal. Generic indications for intubation include need for protection of airway, need for ventilatory support, + oxygenation support. Intubation is to be performed en route except where it is impossible to ventilate the patient. Inability to ventilate requires immediate intubation prior to transport. A list of specific indications and the procedure for a surgical airway exists in the protocols. See Study Binder for more detail.

### Breathing: Questions 3 & 4
The BCAS protocols for breathing intervention mandate an increase in oxygen flow or delivery method to achieve SAO2<95%, including assisted ventilation if required. Oxygen is mandated for any "moderate to severely traumatized patient who may be bleeding". The indications and procedures for a needle thoracentesis exist in the protocols. High flow oxygen and ventilation assistance is advised for major burns. See Study Binder for more detail.

### Circulation: Questions 5 & 6
The BCAS protocols for hypovolemia advise 500 ml N/S boluses for BP<90 mm Hg to a maximum volume of 2000 ml. Mechanism of injury should be considered when determining whether patients have "hypovolemia or potential for hypovolemic shock". The protocol may be used in patients with BP>90 if "shock is anticipated because of the mechanism of injury, the nature and extent of injuries, or the patient's condition". An attendant may attempt an IV a maximum of 3 times. The procedure for external jugular cannulation exists in the protocols. The rule of 9's is mandated for burns > 20% BSA. See Study Binder for more detail.

### Cervical Spine: Questions 7 & 8
The BCAS protocols state that stabilization of the cervical spine is indicated in "patients where spinal injury is suspected or likely because of the mechanism or nature of injury, or unconscious patients where trauma cannot be reasonably ruled out". See Study Binder for more detail.

### Extremity Injuries: Questions 9 - 11
The BCAS protocols state that fracture management is indicated in "suspected limb or joint fractures, dislocations, severe sprains". Assessment of distal neurovascular status is mandated before and after realignment, manipulation or splinting. The indications and procedures for applying a Sagar (traction) splint exist in the protocols. Entonox is written as a BLS and ALS intervention for "pain" along with a list of contraindications and cautions. ALS attendants may consider Morphine Sulphate is an option for pain management (delayed order). See Study Binder for more detail.

### Code 99: Questions 12 & 13
The Trauma 99 protocol identifies 5 trauma receiving facilities (including VGH) to which patients meeting physiologic or mechanism criteria should be transported, even if this involves bypassing other facilities (provided the trauma receiving facility can be reached "in approximately 20 minutes"). See Study Binder for more detail.
Paramedic Instrument

The same 24 questions were proposed for the paramedic version of the data abstracting form. Fourteen of these questions were included in the final version of the paramedic data abstracting form. Table 4.2 displays the questions and the CVR for each.
Table 4.2 All Proposed Questions and Content Validity Ratios (CVR) Calculated from Expert Panel Review for Paramedic Version of Data Abstracting Form

<table>
<thead>
<tr>
<th>Paramedic Instrument</th>
<th>Question</th>
<th>CVR Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Airway</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was there a need to manage this patient’s airway in the prehospital setting?</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Indicate reason for need to manage airway: Respiratory Failure; Failure of gas exchange; Airway Patency; Airway Protection; Other indications</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>How was this patient’s airway managed? Manually; Oral Airway; Endotracheal Intubation; Cricothyrostomy; Other</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>The paramedic’s management of this patient’s airway was: Optimal; Acceptable; Not Appropriate; Cannot tell</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td><strong>Respiration</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was there a need to manage this patient’s breathing in the prehospital setting?</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Indicate reason for need to manage breathing: Respiratory distress; Oxygenation; Other Indication</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>How was this patient’s breathing managed? Oxygen ____ lpm; BVM without intubation; BVM with intubation; Needle decompression; Other</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>The paramedic’s management of this patient’s breathing was: Optimal; Acceptable; Not Appropriate; Cannot tell</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td><strong>Circulation</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was there a need to manage this patient’s blood pressure in the prehospital setting?</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Indicate reason for need to manage this patient’s blood pressure: Hypovolemia; Systemic BP &lt; 90; Signs of shock; Burns to &gt; 20% BSA; Other Indication</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>How was this patient’s blood pressure managed? IV Solutions; Number of IV Cannula; Amount of fluid; Crew not IV endorsed</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>The paramedic’s management of this patient’s blood pressure was: Compliant with BCAS Protocol; Not compliant with BCAS Protocol</td>
<td>1.00</td>
</tr>
</tbody>
</table>
### Cervical Spine

<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there a need to manage this patient's cervical spine in the prehospital setting?</td>
<td>1.00</td>
</tr>
<tr>
<td>Indicate reason for need to manage patient's cervical spine: LOC; Mechanism of Injury; Injury to head/neck; Associated Injuries; Other indication</td>
<td>0.43</td>
</tr>
<tr>
<td>How was this patient's cervical spine managed? Cervical Collar; Backboard; Clamshell; Other</td>
<td>0.71</td>
</tr>
<tr>
<td>The paramedic's management of this patient's cervical spine was: Optimal; Acceptable; Not appropriate; Cannot tell</td>
<td>1.00</td>
</tr>
</tbody>
</table>

### Other Orthopedic Injuries

<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence of injuries to patient's extremities? Yes; No</td>
<td>0.71</td>
</tr>
<tr>
<td>Was there a need to manage this patient's extremity injuries in the prehospital setting?</td>
<td>1.00</td>
</tr>
<tr>
<td>How was this patient's extremity injuries managed? Splint; Traction splint</td>
<td>0.43</td>
</tr>
<tr>
<td>The paramedic's immobilization of this patient's extremity injuries was: Optimal; Acceptable; Not Appropriate; Cannot tell</td>
<td>1.00</td>
</tr>
<tr>
<td>How was this patient's pain managed: Morphine; Entonox; Pain not managed</td>
<td>0.71</td>
</tr>
<tr>
<td>The paramedic's management of this patient's extremity injuries was: Compliant with BCAS protocol; Not compliant with BCAS protocol</td>
<td>1.00</td>
</tr>
</tbody>
</table>

### Code 99

<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>The patient required diversion to a trauma facility based on: MOI; Anatomical findings; Physiologic presentation</td>
<td>1.00</td>
</tr>
<tr>
<td>The paramedic Comply; Did not Comply; with Code 99 Protocol</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The final version of the paramedic form is displayed as figure 4.2.
### Figure 4.2 Final Version of Paramedic Data Abstracting Form

<table>
<thead>
<tr>
<th><strong>Airway</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was there a need to intervene to manage this patient’s airway in the prehospital setting?</td>
</tr>
<tr>
<td>1 _ Yes  2 _ No</td>
</tr>
<tr>
<td>2. The paramedic’s management of this patient’s airway was:</td>
</tr>
<tr>
<td>1 _ Optimal  4 _ Cannot tell  2 _ Acceptable  5 _ N/A  3 _ Not Appropriate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Breathing</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Was there a need to intervene to manage this patient’s breathing in the prehospital setting?</td>
</tr>
<tr>
<td>1 _ Yes  2 _ No</td>
</tr>
<tr>
<td>4a. How was the patient’s breathing managed?</td>
</tr>
<tr>
<td>1 _ Oxygen ___ lpm  2 _ BVM without Intubation  3 _ BVM with Intubation  4 _ Needle Decompression  5 _ Other  6 _ Cannot Tell</td>
</tr>
<tr>
<td>4. The paramedic’s management of this patient’s breathing was:</td>
</tr>
<tr>
<td>1 _ Optimal  4 _ Cannot tell  2 _ Acceptable  5 _ N/A  3 _ Not Appropriate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Circulation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Was there a need to intervene to manage this patient’s blood pressure in the prehospital setting?</td>
</tr>
<tr>
<td>1 _ Yes  2 _ No</td>
</tr>
<tr>
<td>6. The paramedic’s management of this patient’s blood pressure was:</td>
</tr>
<tr>
<td>1 _ Compliant with BCAS Hypovolemia Protocol  2 _ Not Compliant with BCAS Hypovolemia Protocol  3 _ N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cervical Spine</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Was there a need to intervene to manage this patient’s cervical spine in the prehospital setting?</td>
</tr>
<tr>
<td>1 _ Yes  2 _ No</td>
</tr>
<tr>
<td>8. The paramedic’s management of this patient’s cervical spine was:</td>
</tr>
<tr>
<td>1 _ Optimal  4 _ Cannot tell  2 _ Acceptable  5 _ N/A  3 _ Not Appropriate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Extremity Injuries</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Was there a need to intervene to manage this patient’s extremity injuries in the prehospital setting?</td>
</tr>
<tr>
<td>1 _ Yes  2 _ No</td>
</tr>
<tr>
<td>10. The paramedic’s immobilization of this patient’s extremity injuries was:</td>
</tr>
<tr>
<td>1 _ Optimal  4 _ Cannot tell  2 _ Acceptable  5 _ N/A  3 _ Not Appropriate</td>
</tr>
<tr>
<td>11. The paramedic’s management of this patient’s extremity injuries was:</td>
</tr>
<tr>
<td>1 _ Compliant with BCAS Pain Management Protocol  2 _ Not Compliant with BCAS Pain Management Protocol  3 _ N/A  4 _ Cannot tell</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Code 99</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>12. The patient required diversion to a trauma facility based on:</td>
</tr>
<tr>
<td>1 _ Mechanism of injury  2 _ Physiologic Presentation</td>
</tr>
</tbody>
</table>
The paramedic was:
1. Compliant with Code 99 Protocol
2. Not Compliant with Code 99 Protocol
3. N/A
4. Cannot tell

Guide to Questions

Airway: Questions 1 & 2
The BCAS protocols for airway intervention are minimal. Generic indications for intubation include need for protection of airway, need for ventilatory support, + oxygenation support. Intubation is to be performed en route except where it is impossible to ventilate the patient. Inability to ventilate requires immediate intubation prior to transport. A list of specific indications and the procedure for a surgical airway exists in the protocols. See Study Binder for more detail.

Breathing: Questions 3 & 4
The BCAS protocols for breathing intervention mandate an increase in oxygen flow or delivery method to achieve SAO2<95%, including assisted ventilation if required. Oxygen is mandated for any "moderate to severely traumatized patient who may be bleeding". The indications and procedures for a needle thoracentesis exist in the protocols. High flow oxygen and ventilation assistance is advised for major burns. See Study Binder for more detail.

Circulation: Questions 5 & 6
The BCAS protocols for hypovolemia advise 500 ml N/S boluses for BP<90 mm Hg to a maximum volume of 2000 ml. Mechanism of injury should be considered when determining whether patients have "hypovolemia or potential for hypovolemic shock". The protocol may be used in patients with BP>90 if "shock is anticipated because of the mechanism of injury, the nature and extent of injuries, or the patient's condition". An attendant may attempt an IV a maximum of 3 times. The procedure for external jugular cannulation exists in the protocols. The rule of 9's is mandated for burns >20% BSA. See Study Binder for more detail.

Cervical Spine: Questions 7 & 8
The BCAS protocols state that stabilization of the cervical spine is indicated in "patients where spinal injury is suspected or likely because of the mechanism or nature of injury, or unconscious patients where trauma cannot be reasonably ruled out". See Study Binder for more detail.

Extremity Injuries: Questions 9 - 11
The BCAS protocols state that fracture management is indicated in "suspected limb or joint fractures, dislocations, severe sprains". Assessment of distal neurovascular status is mandated before and after realignment, manipulation or splinting. The indications and procedures for applying a Sagar (traction) splint exist in the protocols. Entonox is written as a BLS and ALS intervention for "pain" along with a list of contraindications and cautions. ALS attendants may consider Morphine Sulphate is an option for pain management (delayed order). See Study Binder for more detail.

Code 99: Questions 12 & 13
The Trauma 99 protocol identifies 5 trauma receiving facilities (including VGH) to which patients meeting physiologic or mechanism criteria should be transported, even if this involves bypassing other facilities (provided the trauma receiving facility can be reached "in approximately 20 minutes"). See Study Binder for more detail.
Case Enrolment

A total of 206 trauma patients arrived and were initially screened by the EP or research assistant for eligibility for inclusion in the study. Of these 88 were excluded from the study. Figure 4.3 displays the enrollment and exclusion of these trauma patients with reasons for exclusion. Thirty-four of these cases were ineligible because they did not involve trauma or they were secondary transfers from another health care facility. Twelve cases were missing the BCAS patient care record. In 38 of the eligible cases, EPs failed to complete a form within 48 hours, and in four of the cases, both the BCAS patient care report, and the EP data abstracting form were missing, totaling 88 cases excluded from the study.
There were no statistically significant differences between the included group and the excluded group based on paramedic documentation of age, gender, systolic blood pressure, or Glasgow Coma Scale. See table 4.3 for the distribution of these parameters by groups.

Table 4.3 Inclusion/Exclusion Traits of Study Cases

<table>
<thead>
<tr>
<th></th>
<th>Included</th>
<th>Excluded</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Age</td>
<td>38.1</td>
<td>35.7</td>
<td>0.33</td>
</tr>
<tr>
<td>Gender (% males)</td>
<td>69</td>
<td>65</td>
<td>0.31</td>
</tr>
<tr>
<td>Average Systolic BP</td>
<td>125</td>
<td>120</td>
<td>0.34</td>
</tr>
<tr>
<td>Average GCS score</td>
<td>13.13</td>
<td>12.76</td>
<td>0.57</td>
</tr>
</tbody>
</table>

36
Sixty cases were selected from the Trauma Registry based on the selection criteria outlined above. Of these, the unique identifier number used to track the original BCAS patient care report was incorrect for eight cases. For two other cases the quality of the photocopy provided by BCAS records management was of poor quality. In total there were fifty cases enrolled from the trauma registry. A total of 168 cases were available to evaluate inter-rater reliability; 118 for validity, and 50 for intra-rater reliability.

Quantifying Reliability and Validity

**Inter-rater Reliability**

There were two types of questions included in both data collection forms. The first are indicator questions. These are questions that ask whether there was a need to intervene to management the patient in the prehospital environment. The second type is evaluative. These ask the rater to determine if the indicated protocol was complied with or if the treatments provided were appropriate. A third type of question only appears in the paramedic instrument. This question is simply an information question, wherein the rater is asked to identify the type of respiratory management that was performed by the paramedic. Of the 14 questions included in the paramedic instrument, only 3 had a kappa statistic exceeding 0.7. Two of these were indicator questions:

Question 1. “Was there a need to intervene to manage this patient’s airway in the prehospital setting?” Kappa statistic = 0.83 (95% CI = 0.823 to 0.84); and

Question 7. “Was there a need to intervene to manage this patient’s cervical spine in the prehospital setting?” Kappa statistic = 0.83 (95% CI = 0.82 to 0.83).

The third question was the only information question appearing on either form.
Question 4 a. "How was this patient’s breathing managed?" Kappa statistic = 0.79 (95% CI = 0.79 to 0.80).

The mean kappa statistic for all questions included in the data collection form was 0.33 (95% CI = 0.34 to 0.34) with a range of -0.09 to 0.83. All evaluative questions had kappa statistics below 0.3. Table 2.1 displays the kappa statistics per question. All of the evaluative questions (2, 4, 6, 8, 10, 11, and 13) had mean kappa statistics below 0.25, with the highest recorded for Breathing (kappa = 0.23, 95% CI 0.22 – 0.23). The three explicit protocol questions (8, 11 and 13) had the lowest kappa statistics: -0.06, -0.08, and -0.03 respectively.
Kappa statistics were also calculated for each question with “cannot tell” and “N/A” included. When “cannot tell” and “N/A” are excluded from the analysis, the raters must agree that the question is relevant and there is sufficient information to answer the question before the question is included in the analysis. This form of measurement tends to focus the agreement analysis on the content of the question. When “cannot tell” and “N/A” are included in the analysis, the raters need not agree on relevance or sufficiency of information in the material being rated. This tends to measure how the raters view the quality of documentation as well as the content of the question. Including “cannot tell” and “N/A” permits an indirect measurement of agreement on the quality of the documentation.

When “cannot tell” and “N/A” responses were included, the number and specific questions with kappa statistics over 0.70 remained the same, but the overall mean kappa statistic for all questions improved to 0.42 (95% CI = 0.41 to 0.42) with a range of 0.13 to 0.96.
0.83. Figure 4.5 displays the mean kappa statistics per question when “cannot tell” and “N/A” are included.

**Figure 4.5 Mean Kappa Statistics and 95% Confidence Interval per Question with ‘cannot tell’ and ‘N/A’ included**

With “cannot tell” and “N/A” included, the evaluation questions still failed to achieve the 0.70 cut-off. However, there was a less clear distinction between the explicit protocol questions and questions asking about appropriateness of treatments. The lowest kappa statistic was calculated for Question 13, the question asking if the trauma diversion protocol had been complied with. The other explicit protocol questions, eight and eleven, had mean kappa statistics of 0.32 and 0.24.

Overall mean kappa statistics were calculated for groups of raters. Groupings were based on crew qualification (Basic Life Support reviewers only = BLS, Advanced Life Support reviewers only = ALS, and Quality Improvement Coordinator reviewers only = QI). BLS reviewers had better agreement with each other (kappa statistic=0.47) compared to either ALS reviewers (kappa statistic = 0.34) or QI reviews (kappa statistic...
= 0.35). Table 4.4 displays the kappa statistics and 95% confidence intervals for the entire group of raters and for each qualification group.

Table 4.4 Mean Kappa Statistics and Confidence Intervals (95%) for Raters Grouped by Qualification

<table>
<thead>
<tr>
<th>Rater Group</th>
<th>LCI</th>
<th>Mean Kappa</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Raters</td>
<td>0.33</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td>BLS only</td>
<td>0.46</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>ALS Only</td>
<td>0.33</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>QI Only</td>
<td>0.34</td>
<td>0.35</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only.

LCI Refers to the lower interval of the 95% confidence limit of the mean. UCI refers to the upper interval of the 95% confidence limit of the mean.

When “cannot tell” and “N/A” were included for the calculation of the kappa statistics, the QI reviewers had the highest level of agreement (kappa statistic = 0.44) compared to either the ALS reviewers (kappa statistic = 0.40) or the BLS reviewers (kappa statistic = 0.39). Table 4.5 displays kappa statistics for the entire group and for each qualification group with “cannot tell” and “N/A” values included. Keeping in mind that including “cannot tell” and “N/A” is intended to measure the amount of agreement over the quality of documentation as well as the treatments provided, the change in order between the QI only raters and the BLS only raters should be noted.

Table 4.5 Mean Kappa Statistics and Confidence Intervals (95%) for Raters Grouped by Qualification. “cannot tell” and “N/A” included.

<table>
<thead>
<tr>
<th>Rater Group</th>
<th>LCI</th>
<th>Mean Kappa</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Raters</td>
<td>0.41</td>
<td>0.42</td>
<td>0.42</td>
</tr>
<tr>
<td>BLS only</td>
<td>0.38</td>
<td>0.38</td>
<td>0.39</td>
</tr>
<tr>
<td>ALS Only</td>
<td>0.40</td>
<td>0.40</td>
<td>0.41</td>
</tr>
<tr>
<td>QI Only</td>
<td>0.43</td>
<td>0.44</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only.

LCI Refers to the lower interval of the 95% confidence limit of the mean. UCI refers to the upper interval of the 95% confidence limit of the mean.
In an attempt to determine if any single reviewer was an outlier, causing the grouped kappa statistic to be artificially reduced, the reviewers were systematically removed from the rater group one at a time. Removing raters 3 (BLS), 4 (ALS) and 6 (BLS) resulted in an improvement in the kappa statistic compared to the entire group. Removing rater 4 resulted in the greatest improvement (kappa statistic = 0.36). Figure 4.5 displays the kappa statistics and 95% confidence intervals for the group of raters and for the group with each rater systematically removed.

**Figure 4.6 Mean Kappa Statistic and 95% Confidence Intervals for the Rater Group, with Individual Raters Systematically Removed.**

Note: Rater One = ALS, Rater Two = QI Coordinator, Rater Three = BLS, Rater Four = ALS, Rater 5 = QI Coordinator, and Rater 6 = BLS.

When “cannot tell” and “N/A” are included and each rater is systematically removed from the group, removing raters 3 and 4 resulted in an improvement of the kappa statistic. Removing rater 3 (BLS) resulted in the greatest improvement in the kappa statistic (kappa statistic = 0.45). Figure 4.6 displays kappa statistics and 95%
confidence intervals for the group of raters and for the group with each rater systematically removed.

**Figure 4.7** Mean Kappa Statistic and 95% Confidence Intervals for the Rater Group, with Individual Raters Systematically Removed. “cannot tell” and “N/A included.

---

**Intra-Rater Reliability**

Fifty cases were rated twice by all the raters after a 3-month washout period. The 50 cases all came from the Trauma Registry. Refer to Figure 3.1 on page 24. Intra-rater reliability is calculated as the agreement between two ratings made by the same person. A three-month washout period was used to reduce the possibility of recall bias, which would have spuriously inflated the kappa statistics.

Six questions (1, 4a, 7, 9, 10, and 11) exceeded the threshold kappa statistic of 0.70. Most questions were indicator questions (1, 7, and 9). In addition, kappa statistics in the extremity injury dimension (question 9, 10, and 11) exceeded the 0.70 threshold.
Table 4.7 displays per question mean kappa statistics and 95% confidence intervals for intra-rater reliability.

**Figure 4.8 Intra-Rater Reliability  Mean Kappa Statistics and 95% Confidence Intervals Per Question**

Overall, agreement between the two ratings for all questions and all raters was good with a kappa statistic of 0.62 (95% CI 0.62 to 0.63). The range for all questions by rater was 0.71 (95% CI 0.70 to 0.72) for rater 1 (ALS) to 0.41 (95% CI 0.39 to 0.42) for rater 6 (BLS). Table 4.6 shows the mean kappa statistics and confidence limits for all raters.
Table 4.6 Intra-Rater Reliability Mean Kappa Statistics and 95% Confidence Intervals for Individual Raters

<table>
<thead>
<tr>
<th>Rater</th>
<th>LCI</th>
<th>Mean</th>
<th>Kappa</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td>0.70</td>
<td>0.71</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Rater 2</td>
<td>0.66</td>
<td>0.67</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td>0.65</td>
<td>0.66</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Rater 4</td>
<td>0.62</td>
<td>0.63</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Rater 5</td>
<td>0.63</td>
<td>0.64</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Rater 6</td>
<td>0.39</td>
<td>0.41</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>All Raters</td>
<td>0.61</td>
<td>0.62</td>
<td>0.63</td>
<td></td>
</tr>
</tbody>
</table>

Note: Rater One = ALS, Rater Two = QI Coordinator, Rater Three = BLS, Rater Four = ALS, Rater 5 = QI Coordinator, and Rater 6 = BLS.

BLS raters displayed the lowest level of agreement with themselves based on repeat ratings, with a mean kappa statistic of 0.53 (95% CI = 0.52 to 0.54), while both ALS raters and QI raters both exceeded the overall mean intra-rater kappa statistic.

Figure 4.7 displays the mean kappa statistics and 95% confidence intervals for all raters and the three qualification groupings.

Table 4.7 Intra-Rater Reliability Mean Kappa Statistics and 95% Confidence Intervals for Raters Grouped by Qualification.

<table>
<thead>
<tr>
<th></th>
<th>LCI</th>
<th>Mean</th>
<th>Kappa</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Raters</td>
<td>0.61</td>
<td>0.62</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>BLS Only</td>
<td>0.52</td>
<td>0.53</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>ALS Only</td>
<td>0.66</td>
<td>0.67</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>QI Only</td>
<td>0.64</td>
<td>0.66</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only.

Removing rater 6 (BLS) improved the mean kappa statistic for the data abstracting tool significantly, while removing other raters did not improve or reduce the mean kappa statistically. Figure 4.8 displays the mean kappa statistics and 95% confidence intervals for the group while systematically removing each rater.
Byrt and colleagues [48] derived a method for quantifying the prevalence and bias in the contingency tables used for deriving kappa statistics. Prevalence and Bias Adjusted Kappa (PABAK) is a kappa statistic, which adjusts the observed and expected agreement based on the prevalence of responses and any bias in the responses. Mckinnon [49] developed a macro for MS Excel which calculates PABAK based on cell distributions. Based on this macro, PABAK was calculated for all binary questions (1, 3, 5, 7, 9, 10, and 13) as a means of assessing if the low level of agreement from the instrument resulted from prevalence or bias in the raters' responses.

Adjusting for prevalence and bias in the intra-rater kappa statistics resulted in a statistically significant improvement in questions 1, 3, 5, 6, 7, 12a (both mechanism and physiology-coded to mechanism only), and 13, but a statistically significant decline in question 11. Figure 4.9 displays both kappa and PABAK and the 95% confidence intervals for the rater group, with individual raters systematically removed.
intervals for all binary questions. The greatest differences between kappa and PABAK were for questions 3, 6, and 13. For question 13, adjustment for prevalence and bias resulted in a measure of agreement exceeding the 0.70 threshold. Question 6, which asked if the prehospital treatments were compliant with the circulatory protocol, displayed an increase to 0.65 with the calculation of PABAK. (Refer to Figure 4.2 on page 42 for a list of the questions and responses.)

**Figure 4.10 Intra-Rater Reliability**  Kappa and PABAK Statistics and 95% Confidence Intervals for all Binary Questions

BLS raters had significantly lower agreement for the entire data collection form when kappa statistic and PABAK were combined. Figure 4.10 plots mean kappa statistics and PABAK for all questions by groups of raters.
Table 4.8 Intra-Rater Reliability Mean Kappa (including PABAK) Statistics and Confidence Intervals (95%) for Raters Grouped by Qualification.

<table>
<thead>
<tr>
<th></th>
<th>Mean Agreement Statistic</th>
<th>LCI</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Rater</td>
<td></td>
<td>0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>BLS Only</td>
<td></td>
<td>0.62</td>
<td>0.63</td>
</tr>
<tr>
<td>ALS Only</td>
<td></td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>QI Only</td>
<td></td>
<td>0.71</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only. LCI Refers to the lower interval of the 95% confidence limit of the mean. UCI refers to the upper interval of the 95% confidence limit of the mean.

When rater 6 (mean agreement statistic = 0.41) was removed from the group, there was a significant improvement in the intra-rater reliability for the entire instrument, resulting in overall agreement meeting the 0.70 threshold. Removing other raters did not significantly affect measurement of intra-rater agreement when PABAK was calculated. Figure 4.10 plots the mean kappa and PABAK statistics for the group intra-rater reliability while systematically removing each rater.
Prospective ratings provided by the receiving EPs were accepted *a priori* as the gold standard measure for this thesis. In total 19 physicians rated 118 trauma cases, with a mean number of cases rated per physician of 6.25 and a range of 1 to 11 cases.

Only two questions had correlations greater than 0.70; these were questions 1 with mean phi of 0.79 (95% CI 0.79 to 0.80) and 7 with a mean phi of 0.79 (95% CI 0.78 to 0.79). Two questions had negative correlation coefficients: question 4 with a mean Spearman’s coefficient of -0.01 (95% CI -0.14 to -0.10); and question 6 with a mean Spearman’s coefficient of -0.12 (95% CI -0.12 to -0.09). In response to question 4, physicians were more likely to rate the paramedic’s management of the patient’s breathing as optimal, while the peer raters rated the management as acceptable. In the case of question 6, physicians were more likely to rate the paramedic’s management of
the patient’s blood pressure as compliant, while the peer raters rated the management as not compliant or not applicable (N/A). In the two cases where the physicians rated the management as not compliant, the peer raters were more likely to rate the management as compliant. Figure 4.11 is a plot of the mean correlation coefficients for each question and the 95% confidence intervals.

Figure 4.12 Mean Correlation and 95% Confidence Intervals between Prospective EP Ratings and Retrospective Peer Ratings for Each Question

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean Correlation</th>
<th>LCI</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BLS rater’s ratings had the highest level of correlation with the physician ratings (mean correlation 0.34 with 95% Confidence Intervals of 0.33 to 0.36), while the QI rater’s had the lowest level of correlation with the physician ratings (mean correlation 0.28 with 95% CI of 0.27 to 0.3). Table 4.9 displays the mean correlations and 95% confidence intervals between all raters and the physician raters as grouped by rater qualification.
Table 4.9 Mean Correlation and 95% Confidence Intervals for Rater Groups by Qualification

<table>
<thead>
<tr>
<th></th>
<th>LCI</th>
<th>Mean Correlation</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Raters</td>
<td>0.328</td>
<td>0.343</td>
<td>0.357</td>
</tr>
<tr>
<td>BLS Only</td>
<td>0.363</td>
<td>0.378</td>
<td>0.392</td>
</tr>
<tr>
<td>ALS Only</td>
<td>0.325</td>
<td>0.340</td>
<td>0.355</td>
</tr>
<tr>
<td>QI Only</td>
<td>0.268</td>
<td>0.284</td>
<td>0.299</td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only.

LCI Refers to the lower interval of the 95% confidence limit of the mean. UCI refers to the upper interval of the 95% confidence limit of the mean.

Systematic removal of each rater from the group of raters had no statistically significant effect on the level of criterion validity as measured by correlation between physician and peer raters, although removing raters three and six resulted in a non-significant reduction in the correlations. Figure 4.12 plots the mean correlations and 95% confidence intervals for the group of raters with each rater systematically removed from the group.
Figure 4.13 Mean Correlation and 95% Confidence Intervals between Prospective Ratings by EPs and Retrospective Ratings by Peers

Two small sub analyses were undertaken to better understand the low level of correlation between the physician and peer ratings. The first analysis involved identifying which EPs rated which patient. This allowed correlations to be calculated for each EP individually, rather than the EPs as a group. The purpose of this analysis was to evaluate if the EPs who acted as the gold standard measure rated the cases consistently with each other. As the EPs did not rate all of the cases, the level of consistency between the EPs is difficult to measure. However, this analysis did demonstrate some variability among the EPs as observed by the level of agreement between a single EP and the peer rater compared to that observed between the group of EPs and the peer raters. Figure
4.13 displays a plot of the ratings by one randomly selected EP from the group of VGH EPs who rated more cases than the mean number of cases per physicians.\(^2\)

**Figure 4.14 Mean Correlation between a Single Randomly Selected EP (from the Group of VGH EPs) and Peer Raters**

Of the 10 questions that had correlations calculated between this individual physician and the peer raters, 5 had mean correlation coefficients greater than 0.70. Four of these questions were indicator questions (questions 1, 5, 7, and 12a), while one of them was an evaluative question (question 2; "The paramedic's management of this patient's airway was: optimal, acceptable, not appropriate, cannot tell, N/A").

ALS raters had the highest correlation with the individual EP, while there was no difference between the QI raters or BLS raters. Table 4.10 displays a plot of the mean

\(^2\) Due to the low number of cases correlations could not be calculated for all questions. Of the 14 question (including the two different coding schemes for question 12) only 10 had correlations calculated for them.
correlations and 95% confidence intervals between the groups of raters and the individual physician.

**Table 4.10 Mean Correlation and 95% Confidence Intervals between Single Randomly Selected EP and Peer Raters by Qualification Group**

<table>
<thead>
<tr>
<th>Qualification Group</th>
<th>Mean Correlation</th>
<th>LCI</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS Only</td>
<td>0.60</td>
<td>0.70</td>
<td>0.78</td>
</tr>
<tr>
<td>ALS Only</td>
<td>0.91</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td>QI Only</td>
<td>0.63</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>All Raters</td>
<td>0.89</td>
<td>0.92</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only. LCI Refers to the lower interval of the 95% confidence limit of the mean. UCI refers to the upper interval of the 95% confidence limit of the mean.

Excluding rater 1 (ALS) resulted in a significantly lower correlation between the single EP and the rater group; however systematic removal of the other raters did not result in a significant improvement in the correlation between the EP ratings and the peer ratings. Figure 4.14 plots the mean correlation and 95% confidence intervals between the single randomly selected EP and the peer reviewers for all questions with each rater systematically removed from the group.
A second sub analysis involved the creation of a consensus rating between two retrospective raters of a random sample of 19 cases, and the calculation of correlations between these consensus ratings and the retrospective peer ratings.³

Only two questions had correlations over 0.70 when comparing consensus retrospective review and peer review. These were the two cervical spine questions: question 7 (mean correlation 0.91 with 95% CI = 0.89 to 0.93) and question 8 (mean correlation 0.98 with 95% CI = 0.98 to 0.99). This is in contrast to the high level of correlation between EP raters for questions 1 and 7 in the prospective compared to retrospective comparison, and questions 1, 2, 5, 7, and 12a in the individual EP compared to retrospective peer comparison. Figure 4.15 displays a plot of per question correlation

³ Due to small sample size, correlations could not be calculated for questions 2 or 3.
coefficients and 95% confidence intervals between respective consensus review and the peer review.

Figure 4.16 Mean Correlation and 95% Confidence Intervals between Retrospective Consensus Ratings and Peer Raters by Question

![Figure 4.16 Mean Correlation and 95% Confidence Intervals between Retrospective Consensus Ratings and Peer Raters by Question](image)

BLS raters had significantly higher correlation with the retrospective consensus ratings than did any of the other groups or raters. This is in contrast to a significantly lower correlation between BLS raters and the individual EP, but consistent with the findings for the group of EPs. Table 4.10 displays the mean correlations and 95% CI for all questions with the raters grouped by qualification.

Table 4.11 Mean Correlation and 95% Confidence Intervals between Retrospective Consensus Rating and Peer Raters by Qualification Group

<table>
<thead>
<tr>
<th>Qualification</th>
<th>LCI</th>
<th>Mean Correlation</th>
<th>UCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS Only</td>
<td>0.72</td>
<td>0.77</td>
<td>0.81</td>
</tr>
<tr>
<td>ALS Only</td>
<td>0.53</td>
<td>0.60</td>
<td>0.67</td>
</tr>
<tr>
<td>QI Only</td>
<td>0.39</td>
<td>0.48</td>
<td>0.56</td>
</tr>
<tr>
<td>All Raters</td>
<td>0.56</td>
<td>0.64</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Note: BLS refers to Basic Life Support raters only, ALS refers to Advanced Life Support raters only, and QI refers to Quality Improvement Coordinators only.
LCI Refers to the lower interval of the 95% confidence limit of the mean. UCI refers to the upper interval of the 95% confidence limit of the mean.

Systematically removing individual raters from the group of peer raters had no effect on the overall level of correlation between the peer raters and the respective consensus rating. Figure 4.16 displays the mean correlation and 95% confidence intervals for the correlation between the retrospective consensus ratings and peer ratings, with each rater systematically removed from the group.

Figure 4.17 Mean Correlation and 95% Confidence Intervals between Respective Consensus Ratings and Peer Raters
Relationship Between “Cannot Tell” and Missing Responses and Kappa Statistics and Correlations

As no clear relationships had been uncovered in the main analysis, it was decided to calculate the correlation between the responses that indicate poor quality documentation and the kappa statistics used to evaluate inter-rater reliability. There is a moderate negative correlation ($r = 0.49$) between the kappa statistics used to measure inter-rater reliability and the mean number (per rater) of “cannot tell” and missing responses. Figures 4.17 displays the relationship between the mean kappa statistics and the mean number of “cannot tell” and missing responses.

Figure 4.18 Relationship between the Mean Kappa Statistics per Question and the Mean Number of “Cannot Tell” and Missing Responses per Question.

There was a weak negative correlation ($r = 0.13$) between the mean correlation statistics for each question and the mean number of “cannot tell” and missing responses for each
question. Figure 4.18 displays the relationship between the mean correlations and the mean number of "cannot tell" and missing responses for each rater.

**Figure 4.19  Relationship between the Mean Correlation Statistics per Question and the Mean Number of "Cannot Tell" and Missing Responses per Question**

![Scatter Plot of Mean Correlation per Question and Mean Number of Cannot Tell and Missing Responses](image)

The relationship between the use of a "no" response to a question asking if an intervention was required, and an "N/A" response in the corresponding question asking about appropriateness or compliance was also explored. The proportionate responses for the first three dimensions (airway, breathing and circulation) varied substantially between questions and between raters. Table 4.12 displays the combined number of "N/A" responses for the evaluation questions (2, 4, 6) and "no" responses to the indicator questions (1, 3, 5).
Table 4.12 Proportion of “N/A” Responses to Questions 2, 4, and 6 when “No” Was the Indicated Response to Question 1, 3, and 5, by Rater

<table>
<thead>
<tr>
<th>Rater</th>
<th>N/A</th>
<th>No</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td>163</td>
<td>318</td>
<td>51.2%</td>
</tr>
<tr>
<td>Rater 2</td>
<td>145</td>
<td>205</td>
<td>70.7%</td>
</tr>
<tr>
<td>Rater 3</td>
<td>29</td>
<td>166</td>
<td>17.5%</td>
</tr>
<tr>
<td>Rater 4</td>
<td>159</td>
<td>273</td>
<td>58.2%</td>
</tr>
<tr>
<td>Rater 5</td>
<td>165</td>
<td>217</td>
<td>76.0%</td>
</tr>
<tr>
<td>Rater 6</td>
<td>154</td>
<td>220</td>
<td>70.0%</td>
</tr>
<tr>
<td>Total</td>
<td>815</td>
<td>1399</td>
<td>58.3%</td>
</tr>
</tbody>
</table>

Note: Rate 1 = ALS, Rater 2 = QI, Rater 3 = BLS, Rater 4 = ALS, Rater 5 = QI, and Rater 6 = BLS.

The proportion of “N/A” responses that occurred after a “no” response was recorded in the preceding indicator question varied significantly between the raters (Chi Squared statistic of 34.8 with 5 degrees of freedom. P-value < .001) leading to the conclusion that each of the raters interpreted the questions differently. If rater 3 is removed from the Chi Square test, the probability that the raters differ from each other in the use of the responses is still less than 0.05 (Chi Squared statistics 10.6, with 4 degrees of freedom). At the same time, the raters failed to respond to the questions in a manner that was consistent with each other. A chi square test of the proportion of N/A responses following a no response was rejected at the 0.05 level for each of the raters.
5. Discussion

The major findings of this study are twofold. First, despite being the result of considerable efforts to design a content valid data collection tool to measure processes of care by following a method prescribed for such purposes [32], this tool failed to demonstrate satisfactory validity and reliability. The methods used here included a commonly accepted gold standard measure for EMS research, trained QI coordinators, consultation with local EMS experts on two separate occasions, and focus on a patient population that is of considerable interest to EMS systems. However, inter-rater and intra-rater agreement, and correlation with a prospectively collected EP rating all failed to produce mean values of at least 0.70 for all questions.

The second major finding is that there was no single potential source of error among the six that were investigated that accounted for the poor measurement properties of the data collection form. This will be discussed in more detail below, but overall these findings call into question EMS research based on retrospective peer review, and more distressingly, offer no immediate or easy strategies for the problem of poor measurement of process of care in EMS.

This discussion will focus on the potential sources of error that were explored in this thesis, and broaden out to discuss relevant EMS, emergency, and other clinical medical areas where possible. Figure 5.1 displays the proposed sources of measurement error are displayed. The first potential source of error is the protocols that were used as the criteria upon which the patient records were evaluated. The second potential source of error is the medical records themselves. The third potential source of error is the paramedic raters. The fourth potential source of error is the group of physicians who rated
care in the emergency department. The fifth potential source of error is the data abstracting form. And the sixth potential source of error, which is not displayed in figure 5.1, is artifact or bias introduced by the choice of statistics chosen to represent reliability and validity.

**Figure 5.1 Potential Source of Measurement Error Examined in this Thesis**

**Potential Sources of Rating Error**

**Protocols**

The British Columbia Ambulance Service directs paramedics to provide certain treatments through its protocols, which are part of Policy and Procedure[47]. Only three of the trauma treatment dimensions tested here have explicit protocols written: these are the hypovolemia protocol (to manage a patient's blood pressure), the pain protocol (to manage pain in extremity injuries), and Trauma 99 protocol (for diversion of high risk...
patients to trauma receiving centres). The other four dimensions were evaluated based on a combination of guidelines that appear in Policy and Procedure, additional guidelines included in the back of the data abstracting form, and implicit clinical judgment of the peer raters. Brook, McGlynn, and Cleary advocate the adoption of explicit criteria for the evaluation of process of care when best practice has been established [50]. The three dimensions that have explicit protocols are consistent with this principle [51, 52] [53]. However, the expert panel had difficulty articulating explicit criteria upon which the evaluation of airway, respiratory, orthopedic management could be based.

At the same time, even when a medication or procedure may clearly benefit patients there is no guarantee of consistent application of such a treatment. An example is the administration of aspirin in suspected cases of acute myocardial infarction. The proportion of patients who appropriately receive aspirin within the EMS system has been shown to vary substantially even though there is clear evidence of its benefit [54]. Woollard and colleagues concluded, based on these results that the possible reason for such variability might be the variability in the protocols directing prehospital treatments. They emphasize the need, therefore, to create a consistent set of standards that would apply across the entire health system. The findings of this study raise questions about Wollard’s observations and ask whether the variability observed occurred as a result of processes being evaluated or the retrospective reviewers. In either case, explicit protocols, at least within this study, did not guard against poor reliability and validity of retrospective review.

It was hypothesized if any of the dimensions was going to perform well, it would have been those with an explicit protocol in place. In a study, which sought to evaluate
the quality of care for frail elderly patients, Smith and colleagues found that patient outcome measures exhibited better reliability than process of care measures [55]. Indeed, Smith found that process of care measures were unable to achieve a kappa cut-off value of 0.40, which is consistent with the findings of this study. However, Lerner has demonstrated [17] that there is a potentially important triaging process performed by paramedics, which appears to confound outcome measurement in major trauma. So, although patient outcomes may display greater reliability, there is the potential with trauma research that these findings are confounded.

Interestingly, the existence of explicit protocols, as in questions 6, 11, and 13 did not result in either greater reliability or greater criterion validity. Why, then were questions 1 and 7 the only two that performed to an acceptable level of reliability and validity? It would appear that part of the answer is related to a prevalence bias found in both dimensions. In most patient encounters, in EMS, airway management is rarely performed. This is consistent with the findings in this study. The mean number of “no” responses to question 1 was 130.8 out of 168 cases. At the same time, cervical spine management is very frequently performed in the prehospital setting for trauma patients. The mean number of “yes” responses to question 7 was 135 out of 168 cases. Therefore, as a result of the prevalence of the need for cervical spine management, and the lack of need for airway management, and because both events are either very common or very rare, the peer reviewers had a better implicit understanding of when these managements are required. In other words, for both very rare events, and very common events, the paramedics appear to have a good understanding of when and which prehospital management should occur.
Prevalence, however, is not a sufficient condition for valid and reliable evaluation of the need for prehospital intervention. Question 3, which asks if there was a need to intervene to manage the patients breathing, had equally high prevalence in the responses from the raters. The mean number of “yes” responses was 137 out of the 168 cases. However, there was substantial variability between the raters, as the range of “yes” responses was 108 to 166.

For most of the other questions lack of a prevalence bias may be responsible in part for the poor level of validity and reliability. The mean number of “yes” responses for question 5 (need to intervene to manage the patients blood pressure) was 95. The mean number of “yes” responses for question 9 (need to manage the patients extremity injury) was 77.6. Question 12 asked the rater to indicate the reason for diversion to a trauma receiving facility. The mean number of “mechanism” responses was 73.7, the mean number of physiology responses was 24.7, and the mean number of both “mechanism and physiology” responses was 20.8.

Even though the explicit protocol questions did not perform any better than implicit protocol questions, improved explicit criteria are possible options for some of the dimensions. For example, prospective measurement of oxygen saturation of the blood via a pulse oxymetry could be used in all patients. For intubated patients, Helm and colleagues used several blood gas measurements to allow emergency departments to evaluate the quality of prehospital airway and respiratory management [28]. The use of either blood oxygen saturation or more sophisticated laboratory measures of oxygenation would provide more explicit criteria upon which to evaluate both the need for prehospital management and the quality of that management, and as shown by Helm and colleagues,
provide a valid measure of quality. The high prevalence of “yes” responses to question 3 and the poor reliability of question 4, suggest that inclusion of an explicit measure of de-oxygenation (SAO₂) would have clarified the need for and allowed the raters to distinguish between appropriate care and inappropriate care.

In cases where there is no clear evidence upon which to base protocols, clinical evaluation of the need for management and the quality of management are likely to be poor. Salerno, Wrenn and Slovis note that there are substantial misconceptions about the use of intravenous cannulation among paramedics, which result in substantial protocol deviations [22]. The controversies surrounding the need to initiate an intravenous in the prehospital setting has not abated since 1991 when Salerno’s study was conducted. Researchers continue to look for evidence to either support [53] or discredit [56] prehospital fluid resuscitation. Possible sources of confusion about circulatory support may stem from the protocol objective itself. The purpose of the protocol is simply to maintain a systolic blood pressure at greater than 90 mmhg, and an intravenous may be started when the paramedic anticipates hypovolemic shock due to “mechanism of injury, the nature and extent of the injuries, or the patient’s condition” [47]. BCAS paramedics commonly debate among themselves what it means to anticipate hypovolemic shock.
Documentation of Patient Care

The British Columbia Ambulance Service uses a standard form for the recording of prehospital care. The form requires a combination of specific patient care parameters, such as blood pressure, respiratory rate, pulse rate, and narrative explanation of the injuries found. Regardless of the format of the information, peer reviewers may encounter two problems when reviewing medical records to evaluate quality of care. The first difficulty is illegible or near illegible hand writing, while the second is missing information.

In this study, reviewers were not asked to comment on the quality of the documentation; however, they did have the option to indicate whether they felt the treatments were required or non-compliant or inappropriate. For questions 1 and 7, the mean number of "cannot tell" responses was < 1 out of the 168 cases reviewed. The range of mean number of "cannot tell" or missing responses for the other questions was 1.1 to 41.6. Spaite and colleagues also found that approximately 25% of the required data elements were missing on patient care reports [18]. Introduction of quality improvement feedback was successful in reducing the amount of missing data on patient care forms to less than 5%.

The findings of moderate correlation between the "cannot tell" and missing responses and the kappa statistics offer some support for improving at least the reliability of retrospective review via improved patient record keeping. The very weak correlation between "cannot tell" and missing responses and the correlation coefficients suggests that although retrospective review would be more reliable, it would be no more valid. Direct observation by a trained individual offers the most valid means of evaluating the process.
of care[18, 57, 58]. However, few systems can afford the resources necessary to perform direct observation, so methods must be found to improve the validity of retrospective chart review.

The use of consensus judgment is a common method for performing retrospective review [59]. A small sub analysis, which had two reviewers simultaneously review and rate by consensus 19 randomly selected patient care records was performed. There was a statistically significant improvement in the mean correlations derived by comparing direct observation and retrospective review ($r = 0.34$; 95% confidence intervals 0.33 to 0.36), and the consensus retrospective review ($r = 0.64$; 95% confidence intervals 0.56 to 0.70), which is consistent with the finding that the results of evaluation studies depend upon the source of the data used [60][61][62]. Cales and colleagues [20], Joyce and colleagues [13], and O'Connor and colleagues [19] all found that qualitative data resulted in greater discrepancies than quantitative data leading to the suggestion by Cales [20] that multiple observers be used when clinically important information is being solicited from a patient. Maio noted that paramedics often failed to accurately document if a motor vehicle crash victim was impaired by alcohol, while attributing this lack of documentation to the absence of a standard location for the charting of such information [63], which confirms the findings of Spaite et al [18].

Given the finding here that correlation between retrospective peer review and retrospective consensus review are higher than the correlation between prospective direct observation and retrospective peer review, one conclusion that may be drawn is that the improved correlation resulted from a reduction in the integrity of the gold standard. Paramedics themselves view accuracy and integrity of patient care records as a priority
for the evaluation of the performance of the emergency health services system [64]. In a study of the adherence to treatment guidelines in ambulatory care, Hulka and colleagues noted that 50% of the variability in the rating of adherence to treatment guidelines was due to the quality of the patient records [65]. Therefore, efforts to both improve record quality, and ensure that medical records offer a valid measure of the care provided are needed to improve retrospective review.

**Bias in the Peer Reviewers**

This study sought to evaluate bias by calculating mean agreement and correlation statistics by qualification, and systematically removing the individual raters from the group of raters. There were no consistent trends across the two measures of reliability and validity.

Variability in retrospective review has been noted in previous studies of health care performance [62][55]. Smith and colleagues were able to find a consistent difference between nurses and physicians in their study, which they credit to differences in the type of training that each group receives[55]. They suggest that the difference in training may make the raters more aware or more sensitive to certain findings and that these sensitivities are systematic. However, this study suggests no such variability among paramedics. More surprising is the finding that the QI Coordinators were not systematically more reliable or valid than the BLS or ALS raters.

Pointer and colleagues [66] argued that paramedics ignore or use guidelines to suit their own standards when evaluating if a patient needs to be transported to hospital. However, paramedic judgment was also found to be a valuable addition to another triage study which sought to identify which patients most warranted transport to a trauma
receiving centre [67]. What may be at play in both the literature and this study is the role of a previously noted prevalence bias. In Fries and colleagues study, paramedic judgment was used to augment decisions about the sickest patients. Contrarily, Pointer found that the infractions to the guidelines occurred in the least sick or injured patients where paramedics underestimated the extent of the patient’s illness. Both these studies suggest that paramedics do a fairly good job of agreeing on which patients are sick. The difficulty is in situations in which the patients who are not severely ill or injured.

Physicians

The selection of the emergency room physicians as a criterion measure presents three possible sources of error. The first possible source of error is the comparison of two different sources of data: retrospective chart review and prospective direct observation. Second is the difference in professional qualification between the raters, which has been discussed above. The third possible source of error is that the EPs are too heterogeneous a group to act as a criterion measure.

Gill and colleagues [61] found that there was less agreement between retrospective and prospective methods of evaluating the need for urgent care compared to retrospective and retrospective review. Their results were robust to the specialty of the reviewer. Spaite and colleagues [58] evaluated a method for collecting prehospital care data with the use of direct observation, which they then used to evaluate deficiencies in patient care documentation. EMS systems have used audio tapes [68], and advocate the use of direct infield observation [69]. The improved correlation between the retrospective consensus review support these findings.
It appears from this study that EPs may evaluate compliance and appropriateness of trauma care differently among themselves. When one randomly selected EP was used as the gold standard measure, the correlation between the raters and the EP improved significantly from 0.34 (95% CI 0.33 to 0.36) to 0.92 (95% CI 0.88 to 0.94). This occurred despite extensive efforts to educate the VGH EPs and the availability of all of the BCAS Protocols.

Data Abstracting Form

The items that appear in the data abstracting form may not match very closely with the protocol being evaluated or the treatment construct of the reviewers. This became apparent during the consensus review conducted by me and one of the committee members (RBA). Question 3, “Was there a need to intervene to manage this patient’s breathing in the prehospital setting?” produced several questions. What is an intervention? Is giving the patient supplemental oxygen by simple face mask considered an intervention? What constitutes a need? If the patient does not have any obvious injuries, no physiologic derangement, and they are considered a major trauma based solely on mechanism of injury, is supplemental oxygen a need? Question 5 offered similar difficulties. Is the initiating of intravenous therapy considered an intervention? If the patient is normotensive and does not have any apparent injuries, should an intravenous be initiated?

The guide printed for raters on the opposite face of the data abstracting form may provide some guidance for the raters in the case of question 3. The guide reads “BCAS protocols for breathing intervention mandate an increase in oxygen flow or delivery method to achieve SAO2 < 95%, including assisted vents if required.” However, SAO2
readings were not available on many forms, nor are paramedics directed to monitor
SAO2 on an ongoing basis. Further, the current patient care record lacks a standard
location for the recording of these data, a factor confirmed by Maio [63]. Although
paramedics are required to monitor all vital signs, it is not clear if this actually happens.
Several studies have found that paramedics fail to record vital signs that they do measure,
and that they occasionally fail to measure them [18, 57, 58].

There was also the issue of the negative answers to the indicator question and its
affect on the evaluative question. For example, if the rater indicated that the intervention
was not required, and it was not performed, was this acceptable care for the patient?
Ideally, the rater would have indicated "N/A" when they had judged the intervention as
not needed. However, the raters may have been concerned about capturing the concept of
error when completing the questions. Several options would have then presented
themselves. The raters may be interested in judging if there was an error of commission
in the absence of need. The combination for this response would be that the intervention
was not needed and the paramedic's treatments were not appropriate. The raters may
have been interested in evaluating appropriate omissions, in which case a "no" response
to the indicator question would be followed by any of the following responses to the
second question: "optimal", "acceptable", or "compliant".

There was evidence uncovered that the raters recorded combinations of "no" and
"N/A" responses differently. Consistent response choices among the raters, therefore,
may be as a result of different quality constructs, as discussed above.

Another possible response choice problem in the data collection tool was the use
of "optimal" and "appropriate" as two possible positive responses to questions 2, 4, and
10. It is possible that the amount of agreement among the raters was affected by a response bias towards either of the two possible responses. Therefore, the "optimal" and "appropriate" response choices were collapsed into one positive response and kappa statistics were recalculated for some of the questions that used these responses. One statistically significant difference was found between the uncollapsed and collapsed kappa values for question 2 (uncollapsed kappa = 0.18; 95% CI 0.16 to 0.20; collapsed kappa = 0.23; 95% CI 0.21 to 0.26), but the difference between the two kappa statistics was not sufficient to increase agreement above the 0.70 range. When responses for the EP ratings were collapsed in this manner and correlations between the raters and physicians were recalculated for question 4, there was a non-significant difference between the correlation statistics (uncollapsed r = -0.01; 95% CI -.14 to -.1; collapsed r = -.01; 95% CI -0.02 to -0.01). Therefore, although the choice of multiple positive responses for some of the questions may have affected the amount of agreement between the raters and correlation between the raters and the EP, it does not appear to have caused the agreement and correlations to be below the 0.7 threshold level set a priori.
Byrt and colleagues [48] noted that the kappa statistic is affected by the prevalence and the bias distribution of responses, and it has a tendency to report spurious results if an adjustment for both bias and prevalence is not performed. They derived the Prevalence and Bias Adjusted Kappa (PABAK) to adjust for such factors. PABAK is available for the simple 2x2 table, and so for all binary response questions PABAK was calculated (Refer to figure 4.7). Of the 10 questions that had binary responses 7 had a statistically significant improvement (1, 3, 5, 6, 7, 12a, 13) in agreement based on PABAK rather than unweighted kappa, 1 had a statistically significant decline in agreement (11), and two had non-significant improvements (9, 12b). PABAK resulted in one question improving above the 0.70 threshold for agreement, and two other questions improving above 0.65.

In the cases of multiple category responses, there was a similar prevalence and biased distribution problem, but unfortunately no bias or prevalence adjustment is available for anything greater than binary data. Several problems have been noted with the kappa statistic including variability of results depending on the number of categories [70], and a Simpson’s Paradox-like prevalence bias resulting from failure to control for variability in prevalence of the phenomena being observed [41]. However, unweighted kappa statistics still retain the advantage of measuring agreement rather than association [71], and so make the greatest amount of sense in the validation study setting because the investigator is interested primarily in the amount of agreement between the observers, rather than the association between the observations.
Sensitivity and specificity could have been used to calculate criterion validity. However, the presence of multiple category responses would have entailed the calculation of several measures of sensitivity for each question (corresponding with each of the positive responses), thus adding to the complexity of the analysis. For example, if the “optimal” response is very sensitive for some questions while “appropriate” is sensitive for other questions, what choice should be made about response choice?

Overall, there was not any single factor that led to the poor reliability and validity of this instrument. An incremental approach to improving the data abstracting tool would begin with clarifying the protocols and guidelines for treating trauma patients. Common to other areas of clinical medicine, efforts should be made to improve patient care records. This study intentionally avoided any training of the raters and relied on their own interpretation of the questions and treatment protocols and guideline. Therefore standardized training of the raters would also likely improve the reliability of the ratings. Selection of an appropriated criterion measure that is valid is a particularly difficult issue. Direct observation of paramedics has been advocated by Spaite and colleagues [57], and certainly offers the greatest opportunity to collect process of care information. At the same time, the relative rarity of traumatic injuries, and the possible impact of a Hawthorne effect, make the use of direct observation questionable in so far as feasibility and validity.

The selection of peer reviewers based on qualification did not appear to impact the measurement properties of this data collection form. However, the quality of documentation does appear both in the study and in the literature to affect measurement of validity.
Limitations

There are several limitations to this study. First, other strategies could have been used to quantify the validity and reliability of the data collection tool. For example, a generalizability study could have been used based on use of ANOVA, which would have allowed for the evaluation of sources of error in one statistical process rather than the use of several different calculations of kappa, with comparison. This strategy would have been appropriate had the investigator anticipated such profound problems with the tool before commencing the study; however, this was not the case. Substantial problems with agreement and criterion validity were not found until after the kappa statistics and correlation coefficient had been calculated. At the same time, factor analysis could have been used to determine construct validity rather than correlation statistics to calculate criterion validity. The same argument is forwarded to defend criterion validity, namely, that problems were not anticipated at the outset. Construct validity would likely have been more difficult to achieve by using factor analysis. Further, the dimensions that were under consideration would likely not separate out into distinct factors, although the overall constructs of compliance and appropriateness would have. Factor analysis would also have likely required a larger sample of observers or raters, which would have been a difficulty given both the time and financial constraints of this study.

This study was designed solely for the purpose of quantifying the amount of agreement between and within the raters, and the amount of correlation between the retrospective peer review and the prospective criterion ratings. The sub-analyses presented in the results and the discussions are primarily hypothesis generating in nature, and were not powered for hypothesis testing. Where 95% confidence intervals are
provided, and a statistically significant difference is demonstrated, these results are indicative of a statistically significant difference. Where no statistically significant difference has been demonstrated, the amount of type II error in the study has not been quantified; therefore there remains the possibility that where no difference was discovered there was indeed a difference. At the same time, it is recommended that a valid and reliable questionnaire be created first. Once the questionnaire has been demonstrated to be valid and reliable, then hypothesis testing could be carried out to determine if changing reviewer qualification does indeed result in a consistent pattern of lower reliability and validity.

The use of all of the physicians who work in the ED at VGH is also a limitation of the study. As shown, the physician group did not appear to view compliance in a uniform way, and the selection of a smaller group of EPs who are very familiar with the BCAS protocols and Policy and Procedure might have resulted in improved correlations between the criterion measure and the retrospective ratings.

The raters were not a randomly selected group of peers, but were selected for either their position as QI Coordinators or for their skill and expertise as instructors or curriculum developers of the paramedic programs. Except for one of the BLS raters, all of the raters were engaged in instruction or evaluation of paramedic performance as the primary focus of their duties or as an ancillary function. One of the raters was selected after another candidate who was involved with licensing and the BC College of Paramedics declined to be involved. Therefore, this group of raters is not necessarily a random sample of peers within the BCAS, nor is it a specific select group of specially qualified QI or educator paramedics. Therefore, although there is no reason to think that
these raters differ from any other paramedics who would have volunteered to be engaged in this project, there is still a potential that they do differ, and that a random sample of paramedic peers would have performed differently.

These raters were not trained specifically to evaluate these forms. It is likely that if these peers had been trained to use the form and review the medical records, the reliability of the ratings would have improved. The raters were not trained because it was believed that had this program been performed by BCAS that additional training would not have been forthcoming. In fact because BCAS would only have employed a single rater, there would have been no perceived need to train the QI rater, nor any evidence of a problem. Therefore these results do display the validity and reliability of peer based quality improvement ratings based on a non-standardized peer reviewer basis.

Finally, it is possible that had a more severely injured population of patients been included in this project inter-rater reliability and validity would have been improved. This is based on a premise that the raters would agree with each other more often in the evaluation of need with patients who are more severely injured. Because there are few treatment options, in the face of severe injuries, the paramedics would likely agree with each other on what to do.
6. Implications and Policy Recommendations

This study has demonstrated that the use of a content valid data collection tool can result in substantial amounts of information bias, specifically misclassification of either exposure or outcome depending on how this tool is used. If the data collection tool were used to measure compliance with protocols as a covariate in a multivariate analysis of survival from major trauma, the misclassification would be on exposure. If one were interested in compliance with the protocol in trauma as the outcome for a Quality Improvement program, the misclassification would be on outcome. These forms of misclassification would result, provided the misclassification is non-differential, in a bias towards the null and, therefore a finding of no difference among the variables under study.

To ensure that this misclassification does not occur, content validity should be considered as the first step in the development of any questionnaire or data collection tool in health services research. Further work should then be undertaken to ensure that the tool is reliable and meets some higher standard of validity, such as criterion or predictive validity. Once sufficiently high levels of reliability and validity have been established, the tool may then be put into use. Obviously, undertaking a validity study requires a substantial commitment of time, resources and expertise on the part of the health services provider. However, it is proposed that the additional expenditures at the outset will result in reduced costs when policies are instituted based on spurious or biased results generated by a data collection tool with poor psychometric properties.

Poorly documented patient care and poorly crafted treatment guidelines or protocols likely contributed to some extent to the poor performance of this instrument.
Improvements to documentation of patient care are of paramount importance not only for health services and policy decision-making, but also for medico-legal reasons. Poorly crafted treatment guidelines likely result in variability in the application of the protocols as well as the interpretation of patient care. In both cases, care should be taken to ensure that assumptions are made explicit and any ambiguous or confusing text is clarified. Pilot testing protocols or patient care documents would likely lead to a reduction of errors and ambiguity.

At the outset it was assumed that EPs who have achieved Board qualification are equally qualified to evaluate the quality of care provided by paramedics. Tentatively, this does not appear to be the case. If this is due to a lack of training or knowledge of what paramedics are qualified to do, this is easily remedied. In this study, aside from the guide to the data collection tool that was printed on the reverse of this questionnaire, a document containing all of the training and Policy and Procedure materials that paramedics use in the execution of their duties was available for the EPs. An additional orientation may also be of assistance.

Further Research

There are several research questions that have become evident as a result of this work. First, is there a difference between raters based on training and clinical qualification? BCAS employs ALS-trained paramedics to perform their QI duties, as do many other systems. If it is true, as has been hypothesized based on these results, that the level of clinical qualification is of no consequence to the reliability or validity of the ratings provided, the pool of QI candidate would likely double in many systems. If clinical qualification is being used as a proxy for QI training, or training in data
abstractions and standardization of QI data collection methods, this study suggests that clinical qualification likely has no bearing on these characteristics. Appropriate training for QI and other data abstracting staff is of utmost importance to ensure that the data being collected is done so consistently and accurately.

Second, the calculation of criterion validity was complicated by four variables: variation of ratings among a group of EPs, prospective vs. retrospective evaluation, and source of the data (direct observation vs. chart review) and difference in the qualification between the two sets of raters (physicians and paramedics). This problem suggests a need for several investigations. Are EPs consistent in their judgments of the quality of care provided in EMS? If not, what qualities make a good participant in this sort of study, and does retrospective consensus review offer a better method of review than direct observation? Further, if direct observation is the best criterion, is there a difference in the validity of these ratings based on the qualification of the observer. As suggested by Spaite, [57] would the use of direct observation by a non-clinically trained observer provide valid observational data?

It has been suggested earlier in this study that unvalidated measures of the quality of care in medicine may result in substantial amounts of misclassification bias. Depending on the objective of the study, the misclassification would be in either the outcome if the study were attempting to derive the determinants of quality of care, or the exposure if the study was attempting to measure the impact of varying quality of care on patient outcomes. Further analysis and demonstration of this bias is warranted. As QI programs in EMS mature, interest in deriving determinants of quality of care and measures of the effect of such care will become a greater focus. Demonstrating the effect
of such bias on the operational decisions would ensure that EMS agencies measure their systems as validly and precisely as possible.

The measurement of process of care remains an important if poorly-measured characteristic of prehospital care. EMS systems should improve their data collection process and make attempts to improve quality improvement activities by measuring more than just response times, and a limited set of outcomes on cardiac arrest and major trauma patients. Many clinically important questions remain unanswered with equivocal results. The introduction of process of care measures such as measurement of compliance with protocols may help to answer these important questions. However, reliable and valid data collection tools that can be used to do this measurement remain elusive.
Appendix 1

BCAS Trauma Protocols
BC AMBULANCE SERVICE
PATIENT ASSESSMENT MODULE

Trauma Protocol Compliance by BC Ambulance Service Paramedics - Development and Prospective Validation of a Measurement Instrument

Co-Principal Investigators:
Riyad B. Abu-Laban, MD, FRCPC, MHSc
Douglas L. Andrusiek, BA(Hon), EMAII

Co-Investigators:
Devin R. Harris, MD, CCFP(EM)
Dave Harrison, MD, CCFP(EM), FRCPc
Joseph Ip, MD, FRCPc, MSc
Iain MacPhail, MD, CCFP(EM), MHSc
Richard Simons, MB, BCHr, FRCSC, FACS, FRCS
Samual Sheps, MD, MSC, FRCP
Robert Stenstrom, MD, PhD, CCFP(EM)
Tracy Taulu, RN, BSN
Catherina van Beek, RN, BSN
Patient Assessment Module
Airway Protocols and Procedures
3.4.3.2 MAJOR BURNS

EFFECTIVE DATE: MARCH 22, 2001

REVISION NUMBER: 01-15

PROTOCOL:

INDICATIONS

Patients with major burns greater than 20% of body surface area (BSA). Use "Rule of Nines." 

1. Survey, Vital signs, Oxygen Monitor
2. Transport
3. Secondary survey, vital signs en route
4. Entonox
5. Administer N/S according to the following formula:
   \[ (\text{wt [kg]} \times \% \text{burn}) / 4 = \text{mi/hr} \]
6. Contact Emergency Physician for pain control

1. Assist ventilations and consider intubation for acute respiratory failure.
2. For burns, manage according to the guidelines for Major Burns Protocol.
3. Entonox is contraindicated for suspected inhalation injury because an \( O_2 \) concentration of 100% is desirable.
Needle Cricothyrostomy

Indications for Surgical Airway Procedures:

These techniques are only rarely necessary and the potential for misuse and damage is extremely high. Therefore, they should be used only if all the following conditions apply:

- History of acute respiratory arrest secondary to airway obstruction
- Inability to ventilate using bag-valve-mask
- Inability to intubate
- Visualization of an obstructed airway that cannot be cleared by direct removal of a foreign body or by chest thrusts

Procedure

1. Prep the skin and drape.

2. The cricothyroid membrane is palpated just beneath the inferior prominence of the thyroid cartilage. Insert a large-bore 14G over-the-needle catheter through the skin and cricothyroid membrane into the trachea (Figure A). Remove the needle.

3. Tape the catheter in place (Figure B).

4. Another needle may be inserted alongside the first. One permits continuous oxygen flow and the other allows for exhalation.

5. Conversion to cricothyrotomy or tracheostomy should be accomplished as soon as the patient's condition permits.
Errors and Complications:

Great care must be taken in determining the location of the cricothyroid membrane. A needle inadvertently inserted too high may damage the vocal cords. One inserted too low may injure the thyroid isthmus, inducing hemorrhage. Injury to the posterior wall of the trachea or the esophagus may occur when the needle is advanced too far. Pneumothorax or pneumomediastinum may be produced by malpositioned needles, especially in children.

Failure to provide an expiratory vent for a patient receiving continuous flow oxygen may result in carbon dioxide retention. Remember that this approach to airway management is a temporary one only. Failure to arrange for a more stable airway as soon as the patient’s condition allows is an error in management.
Breathing Protocols and Procedures
3.4.3.2 MAJOR BURNS

EFFECTIVE DATE: MARCH 22, 2001

REVISION NUMBER: 01-15

PROTOCOL:

INDICATIONS: Patients with major burns greater than 20% of body surface area (BSA). Use "Rule of Nines."  

Transport

Secondary survey, vital signs en route

Entonox

Administer N/S according to the following formula:

\[
\text{mL/hr} = \frac{\text{wt (kg)} \times \text{% burn}}{4}
\]

Contact Emergency Physician for pain control

1. Assist ventilations and consider intubation for acute respiratory failure.
2. For burns, manage according to the guidelines for Major Burns Protocol.
3. Entonox is contraindicated for suspected inhalation injury because an \( \text{O}_2 \) concentration of 100% is desirable.
Needle Thoracentesis

Indications:

Tension pneumothorax with deteriorating vital signs, markedly decreased cardiac output, profound shock, or cardiac arrest.

Needle thoracentesis should be by delayed protocol only, unless the patient is in cardiac arrest.

Procedure:

1. Identify insertion point, preferably the 2nd intercostal space in the mid-clavicular line. An alternative site is the 4th or 5th intercostal space in the mid-axillary line.

Acceptable and Preferred Locations for Chest Tube Placement

2. Prepare the skin using Povidone-iodine.

3. Using a 14G over-the-needle catheter connected to a 50-mL syringe, enter the skin, directing the needle to pass above the rib (to avoid the intercostal vessels and nerve that traverse along the inferior border of the ribs).

4. Maintain negative pressure on the syringe. When you enter the pleural cavity, the pressure of the tension pneumothorax may force the plunger out of the syringe.

5. Advance the catheter into the pleural space, and remove the syringe and needle. Place your finger over the end of the catheter to prevent an air leak.
6. Prepare a makeshift flutter valve by cutting off a finger from a sterile glove and cutting off the tip. The makeshift flutter valve can be prepared in advance and kept in your kits indefinitely.

7. Place one end of the flutter valve over the catheter, and secure it in place with tape. On expiration, air will escape from the pleural cavity via the catheter flutter valve. On inspiration, the flutter valve is closed to prevent air from entering the pleural cavity.

8. Upon entering the pleural cavity, you may not aspirate anything. This effectively rules out the diagnosis of tension pneumothorax. Remove the syringe and catheter immediately. You may aspirate blood; this effectively establishes the presence of a hemothorax. Although you may attempt to aspirate as much blood as possible, it is probably not worth doing so in the field. Remove syringe and catheter. Do not discard the blood-filled syringe.

Complications:
- Pneumothorax
- Hemothorax
- Pulmonary contusion or laceration
- Cardiac injury
- Hepatic, gastric, or splenic injury from improper positioning

Appendix B: Procedures – Needle Thoracentesis
Circulation Protocols and Procedures
2.5.3.2 EMA 2: HYPOVOLEMIA

EFFECTIVE DATE: MARCH 8, 2002

REVISION NUMBER: 02-05

INDICATIONS:
Patients with any of the following:
- Hypovolemia
- Systolic BP < 90 mm Hg
- Other clinical signs of shock or
- Patients with burns > 20% BSA (second- and third-degree).

CONTRAINDICATIONS:
Peripheral IVs are contraindicated in patients under 12 years.

CAUTIONS:
Shortness of breath:
Before initiating the Hypovolemia protocol, you must have done the following:
- Completed a primary survey
- Initiated transport
- Obtained a baseline set of vital signs

PROTOCOL:

\[
\begin{align*}
\text{BP} & \geq 90 \text{ mm Hg} \\
\text{Administer IV NS at maintenance rate} \\
\text{Continue with assessment and treatment} \\
\text{If BP falls} & < 90 \text{ mm Hg} \\
\text{Treat as unstable, BP} & < 90 \text{ mm Hg} \\
\text{BP} & < 90 \text{ mm Hg} \\
\text{Administer 500 ml NS rapid infusion} \\
\text{Continue with assessment and treatment} \\
\text{BP} & \geq 90 \text{ mm Hg} \\
\text{Administer IV NS at maintenance rate or} \\
\text{enough to keep BP} & \geq 90 \text{ mm Hg} \\
\text{BP} & < 90 \text{ mm Hg} \\
\text{Treat as unstable, BP} & < 90 \text{ mm Hg} \\
\end{align*}
\]

1. Obtain evidence of loss or a significant quantity of blood or body fluids to support diagnosis of hypovolemia.
2. The Hypovolemia protocol may be used for patients with BP ≥ 90 mm Hg if shock is anticipated because of the mechanism of injury, the nature and extent of the injuries, or the patient's condition.
3. After each 500 ml NS, auscultate the lung bases and reassess BP. If signs and symptoms of pulmonary edema are present, stop fluid bolus and administer NS at maintenance rate.
4. After each 500 ml NS, reassess the BP. While the BP ≥ 90 mm Hg, continue administering 500 ml NS.
   - Start a second IV, if appropriate.
   - Maximum 2000 ml without further orders.
   - Contact EP for further orders.

---

1. BCAS Field Operations Policy and Procedure Manual

Page 95
External Jugular Vein Cannulation

Indications

Patients requiring IV access in whom other sites are not available.

Procedure

1. Place the patient in a supine, head-down position to fill the external jugular vein. Turn the patient's head toward the opposite side.

2. Cleanse and anesthetize the skin with Povidone–alcohol.

3. Align the cannula in the direction of the vein, with the point aimed toward the ipsilateral shoulder.

4. Make venipuncture midway between the angle of the jaw and the mid-clavicular line, "tourniqueting" the vein lightly with one finger above the clavicle.
IV Procedures

IV Initiation

Indications

In prehospital care, the primary indications for IV therapy are to:

- Replace fluid and electrolytes due to hypovolemia and burns
- Administer medications

Please refer to the relevant training level for specific indications within a protocol.

Contraindications

- There should be no more than three attempts at insertion by one attendant.
- EMA II attendants should not initiate peripheral IVs in patients under 12 years of age.

Procedure

1. Gather and prepare equipment:
   - Select and inspect the catheter device.
   - Select and inspect the IV solution and administration set.
   - Prime the IV tubing.

2. Choose and prepare an appropriate site.

3. Initiate IV.

4. Collect venous blood samples, if required.

5. Connect IV tubing and infuse solution.

6. Calculate and maintain an appropriate flow rate.

7. Secure the IV.
Maintenance Rate

EMA I and EMA II IV protocols will contain a reference to a maintenance rate. This maintenance rate is approximately 1 gtt/sec. Two common administration sets are used by BCAS: 10 gtt/mL and 60 gtt/mL. To calculate flow rates, the following formula is used:

\[ \text{gtts per minute} = \frac{\text{volume to be infused} \times \text{set rate}}{\text{time in minutes}} \]

Examples:

- To infuse 500 ml normal saline over 12 hours using a macro-drip set (10 gtt/mL):
  \[ \frac{500 \text{ ml} \times 10 \text{ gtt/min.}}{12 \text{ hours} \times 60 \text{ min.}} = \frac{5500}{720} = 7.5 \text{ gtt/min.} \]

- To infuse 25 ml 5% D/W in 60 minutes using a micro-drip set (60 gtt/mL):
  \[ \frac{25 \text{ ml} \times 60 \text{ gtt/min.}}{60 \text{ min.}} = \frac{1500}{60} = 25 \text{ gtt/min.} \]

IV Maintenance

1. Ensure that the appropriate solution is running.
2. Calculate and maintain the appropriate flow rate.
3. Monitor flow rate and amount of solution.
4. Reassess patient condition and IV on a regular basis (i.e., q 5-15 min):
   - Reassess ABCs and injury sites.
   - Reassess vital signs.
   - Inspect IV site, tubing, and solution bag.
   - Observe for complications and take appropriate measures as necessary.
   - Maintain appropriate flow rate.
   - Change solution bag if required.

EMA I and EMA II attendants are not to manage patients in cases where medications or other additives have been introduced to the IV solution. An appropriate medical escort is required in these cases.

Appendix B: Procedures - IV Procedures
3.4.3.2 MAJOR BURNS

EFFECTIVE DATE: MARCH 22, 2001

REVISION NUMBER: 01-15

INDICATIONS

Patients with major burns greater than 20% of body surface area (BSA). Use “Rule of Nines”

- Transport
- Secondary survey, vital signs en route
- Entonox

Administer N/S according to the following formula:

\[(\text{wt (kg)} \times \text{% burn})/4 = mL/hr\]

Contact Emergency Physician for pain control

---

1. Assist ventilations and consider intubation for acute respiratory failure.
2. For burns, manage according to the guidelines for Major Burns Protocol
3. Entonox is contraindicated for suspected Inhalation Injury because an O₂ concentration of 100% is desirable.
Extremity Injuries Protocols and Procedures
2.5.3.1 EMA 2: PAIN USING ENTONOX

EFFECTIVE DATE: MARCH 8, 2002
REVISION NUMBER: 02-04

PROTOCOL:

INDICATIONS:
- Pain.

CONTRAINDICATIONS:
- Inability to ventilate an enclosed treatment area
- Inability to comply with instructions
- Suspected rib fracture or pneumothorax
- Patient has taken nitroglycerin within the last 5 minutes
- Decompression sickness

CAUTIONS:
- Depressant drugs
- Malignant hyperthermia
- COPD
- Distended abdomen
- Shock

Before initiating the Pain Using Entonox protocol, you must have done the following:
- Completed a primary survey
- Investigated the pain complaint, including severity
- Obtained a baseline set of vital signs including oxygen saturation
- Conducted a history and physical examination sufficient to rule out the contraindications for use of Entonox

Explain to patient:
- Entonox is self-administered
- Effects of Entonox
- Possible side effects

Administer Entonox:
- Patient uses until pain is relieved or side effects appear

Monitor and record:
- Start and stop times of Entonox
- Patient response

1 Entonox may be administered to patients with suspected inhalation injuries if O2 saturation is > 90%.
2 Patients should receive high-flow oxygen when Entonox is discontinued.
3 Discontinue if cyanosis develops.
2.4.3.1 PARAMEDIC 1: PAIN USING ENTONOX

**INDICATIONS**

Pain.

**CONTRAINDICATIONS**

- Inability to ventilate an enclosed treatment area.
- Inability to comply with instructions.
- Suspected inhalation injury.
- Suspected air embolism or pneumothorax.
- Patient has taken nitroglycerin within the last 5 minutes.
- Decompression sickness.

**CAUTIONS**

- Depressant drugs.
- Maxillo-facial injuries.
- COPD.
- Distended Abdomen.
- Shock.

Before initiating the Pain Using Entonox protocol, you must have done the following:

- Completed a primary survey.
- Investigated the pain complaint, including severity.
- Obtained a baseline set of vital signs, including oxygen saturation.
- Conducted a history and physical examination sufficient to rule out the contraindications for use.

Explain to patient:

- Entonox is self-administered.
- Effects of Entonox.
- Possible side effects.

Administer Entonox:

- Patient uses until pain is relieved or side effects appear.

Monitor and record:

- Start and stop times of Entonox.
- Patient response.

Patient may be administered to patients with suspected inhalation injuries if O2 saturation is 100%.

Patients should receive high-flow oxygen when Entonox is discontinued.

Discontinue if cyanosis develops.
Sager Splint

Indications:

Suspected lower-limb fractures between the mid-shaft femur and the mid-shaft tibia/fibula.

Procedure

1. Assess distal circulation, sensation, and function.
2. Apply cold, if appropriate.
3. Ensure that patient is supine, with injured leg in line with the body.
4. Place splint beside injured leg.
5. Secure thigh belt.
6. Apply ankle harness above the malleoli.
7. Apply traction:
   - Closed, mid-shaft fractures: 10% of patient’s body weight to a maximum of 10 lbs.
   - Open fractures or joint injuries: Maximum of 5 lbs.
8. Ensure adequate padding.
9. Stabilize limb and splint by applying three elasticized straps.
10. Reassess distal circulation, sensation, and function.
Code 99 Protocol
In the Lower Mainland there are five trauma receiving facilities:

- Vancouver General Hospital
- Royal Columbian Hospital
- St. Paul's Hospital
- Lion's Gate Hospital
- BC Children's Hospital (age 16 and under)

VGH must accept Trauma 99 patients from within its catchment area even when on CCB.

**St. Paul's**

**YES - Penetrating Trauma (excluding head)**

**NO**

- Head Injuries (blunt or penetrating)
- Spinal Trauma
- Blunt Trauma

Note that non-trauma patients with neurologic complaints (e.g. stroke) can be transported to St. Paul's.

**Children's and Women's Health Centre**

BC Women's Hospital (AKA "Grace Hospital") does not have a standard emergency department and does not receive pregnant trauma patients.
BURN Patients

Critical Burn Criteria
- Burns + other major injury (trauma 99)
- Facial or airway burn + inhalation injury
- 2° burn >20% body surface area (BSA)
- 3° burn >10% BSA in adults or 2% BSA in children
- Any 3° burn involving eyes, neck, hands, feet or groin
- Any high voltage electrical burn regardless of size

VGH is the regional burn centre. Therefore:
Critical Burn injury ➔ VGH if <20 min Code 3

If VGH is >20 min:
Transport Code 3 to the closest trauma receiving hospital

If closest trauma hospital >20 min from the scene
Transport to closest Hospital

Note: Patients with non-critical burn injury can be transported to the closest hospital

Pediatric patients who are suspected to have only inhalation injury can be transported to Children’s Hospital if within 15 minutes Code 3.

Nov 2001
Interview Subject: 1.

Interview Date: 18 July 2003

Subject clinical qualification:
MD, FRCPS

Subject duration at clinical qualification:
20 years

Instructions: Read questionnaire item while thinking about reviewing aBCAS crew report. Explain to the interviewer what you believe the question to mean in plain language. Include as much or as little detail as you believe is necessary in providing a comprehensive interpretation of the question. You may also scribe notes on the questionnaire form provided.

Response: Notes: make certain instructions are explicit to BCAS protocol and procedures

Airway Was there a need to manage this patient's airway in the prehospital setting? Yes__ No__

Manage (meaning?) perhaps confusing
Response: Need in what way?
As scientist is necessary to know meaning

Indicate reason for need to manage airway: 
_____ aRespiratory failure
_____ bFailure of gas exchange
_____ Airway Patency _____ Airway Protection _____ Other indications:

Response: Overlap of a & b. b represents bellows failure.
An increasing CO2 or falling O2: is it vent or gas exchange.
I sees as separating hypoxia and ventilatory failure
How was this patient’s airway managed? aManually ___ Oral Airway ___ Endotracheal Intubation ___ Cricothyrostomy ___ Other _______________

Response: think means manual jaw repositioning
          missing suctioning
          add foreign body removal

The paramedic’s management of this patient’s airway was: Optimal ___
          Acceptable ___ Not Appropriate ___ Cannot tell

In relation to training and tools available in the setting?
Response: Therefore will need to know limitations of the paramedics training.

Respiratory
Was there a need to manage this patient’s breathing in the prehospital setting?
    Yes ___ No ___

Response: “manage pt’s breathing” does this mean assist?
          Need clarification

Indicate reason for need to manage breathing:
Respiratory distress: _____ Oxygenation _____
Other Indication ______________

Response: obvious

How was this patient’s breathing managed? Oxygen ___ lpm
          BVM without intubation _____ BVM with intubation _____ Needle decompression ___
Refer to two questions above: So manage does not mean assisting!!
(because oxygen is included)
Response: You have to separate out airway and breathing.

The paramedic’s management of this patient’s breathing was:
Optimal ____ Acceptable ____ Not Appropriate ____ Cannot tell

Response: In relation to skill, training, tools, and setting
Would patient have benefited from fluids in prehospital setting?
Circulation
Was there a need to manage this patient’s blood pressure in the prehospital setting? Yes ____ No ____

Response: “manage this patient’s blood pressure”(?)
do you believe patient would have benefited from... only thing available is fluids
(ongoing thoughts)

Indicate reason for need to manage this patient’s blood pressure:
Hypovolemia ____ Systolic BP < 90 ____ Signs of shock ____
Burns to > 20% BSA ______
Other Indication ______________________________________

Consistent between 1 and 2 for need to manage
Response: Benefit would be obtained from IV fluids because of..... (what JC thinks is the meaning of the question)

How was this patient’s blood pressure managed?
IV Solutions ____
Number of IV Cannula ____
Amount of fluid ____ (N/A) Crew not IV endorsed ______

Response: how was iv fluid managed (does he mean blood pressure?)
The paramedic’s management of this patient’s blood pressure was Compliant ___ Not Compliant ___ with BCAS Hypovolemia Protocol.

Response: Compliant and not compliant sounds punitive

Cervical Spine

Was there a need to manage this patient’s cervical spine in the prehospital setting? Yes ___ No ___

Response: Need is subjective – judgement required

Indicate reason for need to manage patient’s cervical spine:
- LOC ______ Mechanism of Injury ______ Injury to Head/Neck

Associated Injuries ______
Other Indications:

Decreased level of consciousness
Response: Missing neck pain Associated injuries – what does it mean? Needs to be contextually attached

How was this patient’s cervical spine (stabilization) managed? Cervical Collar ______ Backboard ______ Clamshell ______ Other ______

Response: See copy of 1’s notes

The paramedic’s management of this patient’s cervical spine was:
- Optimal _____ Acceptable _____ Not Appropriate _____
- Cannot tell _____

Response: Is this according to stated guidelines?
Other Orthopedic Injuries
Evidence of injuries to patient’s extremities
Yes ___ No ___

Response: Need a verb in statement
Clearer if “potentially unstable orthopedic injuries”

Was there a need to manage this patient’s extremity injuries in the prehospital setting? Yes ___ No ___

Response: Assume bony injuries, not vascular injuries

How was this patient’s extremity injuries managed?
Splint ____ Traction Splint ____ Reduction/Alignment ____

Response: fine

The paramedic’s immobilization of this patient’s extremity injuries was Optimal _____ Acceptable _____ Not Appropriate _____
Cannottell ______

Response: According to protocols or guidelines

How was this patient’s pain managed:
Entonox ____ Morphine ____ Pain not managed ____
Response: See 1's notes Pain by unconscious or other contraindication or patient has no pain Missing patient has no pain Added inappropriate

The paramedic's management of this patient's extremity injuries was Compliant ____ Not Compliant ____ with BCAS pain management protocol.

Prefers consistent with and not consistent with Code 99
The patient required diversion to a trauma facility based on MOI ____ Anatomical findings ____ Physiologic Presentation ____

Response: Asked about guidelines as in when is it okay to not divert

The paramedic Complied ____ Did not Comply ____ with Code 99 Protocol.

Response: Discussed consistent vs complied

Interview Subject: 2.

Interview Date: 18 July 2003

Subject clinical qualification: M.D., FRCPS, FACEP

Subject duration at clinical qualification: 12 years

Instructions: Read questionnaire item while thinking about reviewing a BCAS crew report. Explain to the interviewer what you believe the question to mean in plain language. Include as much or as little detail as you believe is necessary in providing a comprehensive interpretation of the question. You may also scribe notes on the questionnaire form provided.
Airway Was there a need to manage this patient’s airway in the prehospital setting? Yes ___ No ___ What do you mean by manage? Not clear as a stand alone question at first

Response: Believes that “intervene” would be more clear.

Indicate reason for need to manage airway: Respiratory failure _____ Failure of gas exchange _____ Airway Patency _____ Airway Protection _____ Other indications:

Response: No comment/ simple and clear

How was this patient’s airway managed?

aManually _____ Oral Airway _____ Endotracheal Intubation _____ Cricothyrostomy _____ Other ____________________________

Response: foreign body obstruction as well, so need BVM as well include: suctioning, foreign body removal

The paramedic’s management of this patient’s airway was: Optimal _____ Acceptable _____ Not Appropriate _____ Cannot tell
Response: Talked about removing the "paramedic's management" rating
(partial thought process)
Respiratory
Was there a need to manage this patient's breathing in the prehospital setting?
Yes ___ No ___

Response: Prefers intervene
Thinks manage is too global

Indicate reason for need to manage breathing:
Respiratory distress: _____ Oxygenation _____
Other Indication ________________________________

Response: No comment/ simple and clear

How was this patient's breathing managed?
Oxygen ___ lpm
BVM without intubation _____ BVM with intubation _____ Needle decompression _____
Other _____

Response: No comment/ simple and clear

The paramedic's management of this patient's breathing was:
Optimal _____ Acceptable _____ Not Appropriate _____ Cannot tell
Response: No comment/ simple and clear
Circulation Was there a need to manage this patient’s blood pressure in the prehospital setting? Yes ____ No ____

Response: Give fluids for low B.P.

Indicate reason for need to manage this patient’s blood pressure:
- Hypovolemia ____
- Systolic BP < 90 ____
- Signs of shock ____
- Burns to > 20% BSA ____
- Other Indication ____________________________

Response: Only two reasons, what about head trauma?

How was this patient’s blood pressure managed?
- IV Solutions ____
- Number of IV Cannula ____
- Amount of fluid ____
- Crew not IV endorsed ____

Response: No comment/ simple and clear

The paramedic’s management of this patient’s blood pressure was Compliant ____ Not Compliant ____ with BCAS Hypovolemia Protocol.

Response: What about physicians knowing protocols?
Cervical Spine
Was there a need to manage this patient's cervical spine in the prehospital setting?
Yes ☐ No ☐

Response: No comment/ simple and clear

Indicate reason for need to manage patient’s cervical spine: LOC ☐
Mechanism of Injury ☐ Injury to Head/Neck ☐

Associated Injuries ☐
Other Indications:

Response: No comment/ simple and clear

How was this patient’s cervical spine managed?
Cervical Collar ☐
Backboard ☐ Clamshell ☐ Other ☐

Response: No comment/ simple and clear

The paramedic’s management of this patient’s cervical spine was:
Optimal ☐ Acceptable ☐ Not Appropriate ☐ Cannot tell ☐
Response: However – management of BCAS practice not good (meaning over this is done poorly)

Other Orthopedic Injuries Evidence of injuries to patient’s extremities Yes ___ No ___

Response: No comment/ simple and clear

Was there a need to manage this patient’s extremity injuries in the prehospital setting? Yes ___ No ___

Response: No comment/ simple and clear

How was this patient’s extremity injuries managed?
Splint ___ Traction Splint ___ Reduction/Alignment ___

Response: No comment/ simple and clear

The paramedic’s immobilization of this patient’s extremity injuries was
Optimal _____ Acceptable _____ Not Appropriate _____ Cannottell ___

Response: No comment/ simple and clear

Does this get to the severity of the injury?
How was this patient’s pain managed:
Entonox ___ Morphine ___ a Pain not managed ___

Response: a) does this mean no need to manage or pain not managed?

The paramedic’s management of this patient’s extremity injuries was
Compliant ___ Not Compliant ___ with BCAS pain management protocol.

Response: No comment/ simple and clear

Code 99 The patient required diversion to a trauma facility based on MOI
___ Anatomical findings ___ Physiologic Presentation ___

Response: 2 clarifies by saying only two criteria: MOI _____
Physiology _______ (long bone fractures as per physiology)
The paramedic Complied _____ Did not Comply _____ with Code 99 Protocol.

Response: No comment/ simple and clear

Interview Subject: 3

Interview Date:
16 July 2003

Subject clinical qualification: EMA III, ALS 2, Air Evac
Subject duration at clinical qualification: 4 1/2

Instructions: Read questionnaire item while thinking about reviewing aBCAS crew report. Explain to the interviewer what you believe the question to mean in plain language. Include as much or as little detail as you believe is necessary in providing a comprehensive interpretation of the question. You may also scribe notes on the questionnaire form provided.

Airway Was there a need to manage this patient’s airway in the prehospital setting? Yes ___ No ___

Response: Look for indication on form of problem with airway Indications
GCS <9-10, airway complication (vomiting, bleeding), sedatives, c/c eg. Cardiac arrest.

Indicate reason for need to manage airway:
aRespiratory failure ___
bFailure of gas exchange ___
cAirway Patency ___
Protection ___ eOther ___
indications:

Response: use clinical indications as a guide for looking @ form (low oxygen saturation/high or low resp rate) a) notes that mechanical respiration is difficult b) hypoxia indications c) obstructions, GCS, fluid in airway, anaphylaxis, plus diagnosis, burns (notes that b & C are subgroups of each other) an airway you are attempting to improve e) less other indications (such as suctioning not a field)

How was this patient’s airway managed?
aManually ___ bOral Airway ___ cEndotracheal Intubation ___
dCricothyrostomy ___ eOther ________________________________

Response: self evident
using hands, jaw thrust, positioning 3/4 prone
evident- non-definitive assisting device
as defined in good position and confirmed by auscultation and ETCO2
surgical airway
The paramedic's management of this patient's airway was:
Optimal ____
Acceptable ____ Not Appropriate ____ Cannot tell

Response: 3. as the evaluator would decide implicitly on what is optimal – eg optimal in pt with GCS < 6, not optimal for GCS = 14
Optimal most likely to maintain patients condition without undue risk. Requires risk/benefit analysis.
Likely difficult to tell 100%.

Respiratory
Response: (mechanical. Talk about ventilation)
Was there a need to manage this patient's breathing in the prehospital setting? Yes ____ No ____
Response: Indication on form that pt's breathing was inadequate and should be improved. Indications- persistent high or low RR, persistent low SaO2, other signs of hypoxia. Notes outlining respiratory insufficiency. Clues based on diagnosis and treatment. Eg 50 sprays of nitro and 300 of lasix but rate still 50, needs to be treated aggressively.

Indicate reason for need to manage breathing:
Respiratory distress: ____ Oxygenation ____
Other Indication __________________________

Response: RD= more weight on inadequate oxygenation but it depends-Based on form- persistent elevated or low rates difficult as stand alone indications
RD and oxygenation are so intertwined. Times when RD is not related to oxygenation (specifically hypercardic, ketoacidosis) based on PCO2 rather than PO2. Chest injury, pneumothorax.

How was this patient's breathing managed? aOxygen ____ lpm
bBVM without intubation ____ cBVM with intubation ____ dNeedle decompression
____ Other ____
Response: is giving oxygen managing breathing-think of managing mechanical
ventilation rather than oxygenation - oxygen has more to do with oxygenation rather than ventilation
look for note on form for indication look for note on form for indication needle decompression fits better than oxygenation positioning, splinting, control of pain, NG tube/decompress stomach to improve vital capacity

The paramedic's management of this patient's breathing was:
Optimal _____ Acceptable _____ Not Appropriate _____ Cannot tell
Response: Look for indicators of increased ventilation with improved oxygenation.
Airway is check point, but for ventilation you are looking for effectiveness – improvement in hypoxia, improves colour, normalizing of RR, improve LOC, increasing patient comfort

Circulation
Was there a need to manage this patient's blood pressure in the prehospital setting? Yes _____ No _____
BP more objective.
Response: Look at BP - certain ranges - outside need to be managed - below 80 – 90 with other extenuating circ. 210/120 would beg for an intervention or should be considered. Other indicators of poor circ, nailbed colour, LOC, skin colour, skin temp. Because of extenuating circumstances, need to manage something else. Eg failure patient with systolic pressure of 65 mmhg – pt chronic BP, low BP with LOC begs for intervention.

Indicate reason for need to manage this patient's blood pressure:
a) Hypovolemia _____ b) Systolic BP < 90 _____ c) Signs of shock _____
d) Burns to > 20% BSA _____
Other Indication ______________________________
Response: Indicating fluid depleted/circ volume low - assess. A) BP and diagnosis or presentation. Is this patient's low BP suggestive of hypovolemia – is there support. Hypovolemia can be in absence of hypotension- not conjoined – hx vomiting/diarrhea, sepsis, blood on ground, JVP, pulse ox wave form, skin colour and temperature, peripheral presentation and central presentation.
B) self evident
c) as indicated above plus hypotension in presence of c.p., chf, arrhythmia, LOC,
d) notes on form,
e) no other note

How was this patient's blood pressure managed?
a) IV Solutions _____ b) Number of IV Cannula _____ c) Amount of fluid _____ d) Crew
not IV endorsed ______ evidence
Response: ALS – saline lock or infusion also include number and size- attendant
think of large
Other might be appropriate- positioning supine, semi-sitting, trendelburg, trauma
management
The paramedic's management of this patient's blood pressure
was Compliant ___ Not Compliant ___ with BCAS Hypovolemia
Protocol.
Response: Look for protocol and what they did EMA II and PI level is explicit Did
something they should have done or didn't do something they should have.
Distinction between compliant and optimal treatment Perhaps add

Cervical Spine Was there a need
to manage this patient's cervical
spine in the prehospital setting?
Yes ___ No ___
Response: Indicators- risk of c-spine MOI, pain, neuro deficits, standard list for
ruling out (distracting injuries, not intoxicated)
Look for MOI and pt's c-spine not managed – look to see if these factors are ruled
out. Pain, neuro deficits, distracting, intox, GCS < 15, c/c of neck pain with absence of
MOI (and neuro deficits)

Indicate reason for need to manage patient's cervical spine: LOC ______
Mechanism of Injury ______ Injury to Head/Neck

Associated Injuries ______
Other Indications:

Response: Neuro deficits (add)
Reduced LOC vs loss of consciousness, isolated minor injuries
MOI to subjective but okay

How was this patient's cervical spine managed? Cervical Collar ____ Backboard
 ____ Clamshell ____ Other ____
Response: Direct notes or codes Other – sandbags or manual Positioning potentially also - must be supine in code of position

The paramedic's management of this patient's cervical spine was:
Optimal _____ Acceptable _____ Not Appropriate _____
Cannot tell _____

Response: Same as other protocol question
-risk vs. benefit risks are low but there are circumstances when it might present a risk. If they indicate a need to be managed, there must be some specific reason that it was ruled out or managed in a non-traditional way.

Other Orthopedic Injuries Evidence of injuries to patient's extremities Yes _____ No _____
Response: Note regarding injuries
Protocol or treatment that is splinted
Note that a Sagar is used
Patient positioning
mechanism

Was there a need to manage this patient's extremity injuries in the prehospital setting? Yes _____ No _____
Response: Depends on definition of manage – range from position and stabilize to splint and immobilization. *circulatory or nerve impairment as indicator multi-trauma with wrist – ice rather than splint with some stabilization – placed on clamshell

How was this patient's extremity injuries managed?
aSplint _____ bTraction Splint _____ cReduction/Alignment _____
Response: use some adjunct – direct indicator – sam, sagar, clamshell, spine brd
b) c) look for note – neuro or circ which is improved

The paramedic's immobilization of this patient's extremity injuries was
Optimal _____ Acceptable _____ Not Appropriate _____ Cannottell _____
Response: Risk v benefit Delay of transport, stable vs unstable, risk of …. Need to define optimal – we
don’t have ability to optimally manage Is the
optimal contextual or is it global – Must
be based on environment/context

How was this patient’s pain managed:
Entonox _____ Morphine _____ Pain not managed _____

Response: Potential for other _____ as in splinting.
Direct indicators, protocol codes

The paramedic’s management of this patient’s extremity injuries
wasCompliant _____ Not Compliant _____ with BCAS pain
management protocol.
Response: Sticky area when it comes to MS N2O2 is clear – MS is grey for extremities *need clarification of MS guidelines for MS – very loosely laid out
protocol for
pain for ALS

Code 99 The patient required diversion to a trauma facility based on MOI
_____ Anatomical findings _____ Physiologic Presentation _____
Response: Spell out MOI
As per Code 99 protocol
Clearly defined as per protocol

The paramedic Complied _____ Did not Comply _____ with Code 99 Protocol.
Response: Self evident – did pt have indicators and were they taken code 3 to trauma centre within 20 mins
May be complicated by context in the fringe/suburbs
Interview Subject:
Interview Date: 18 July 2003

Subject clinical qualification:

EMA III, ALS II, Air Evac

Subject duration at clinical qualification:

Five Years

Instructions: Read questionnaire item while thinking about reviewing aBCAS crew report. Explain to the interviewer what you believe the question to mean in plain language. Include as much or as little detail as you believe is necessary in providing a comprehensive interpretation of the question. You may also scribble notes on the questionnaire form provided.

Airway Was there a need to manage this patient’s airway in the prehospital setting? Yes ___ No ___

Response: No comment

Indicate reason for need to manage airway: Respiratory failure ___
Failure of gas exchange ___
Airway Patency ___ Airway Protection ___ Other indications:

Response: No comment

Resp arrest means that PCO2 is greater than PO2
Response: Secure airway to deal with resp arrest

How was this patient’s airway managed?
Manually ___ Oral Airway ___ Endotracheal Intubation ___
Cricothyroscopy ___ Other ____________________________
Response: Add: suction, nasal airway, combi-tube

The paramedic's management of this patient's airway was: Optimal _____
Acceptable _____ Not Appropriate _____ Cannot tell

Response: No comment

Respiratory
Was there a need to manage this patient's breathing in the prehospital setting?
Yes _____ No _____

Response: No comment

Indicate reason for need to manage breathing: Respiratory distress: ______
Oxygenation ______
Other Indication ____________________________________________

Response: Thinking of resp arrest or failure.

What about crew qualification field on form- will it be correct?

How was this patient's breathing managed? Oxygen _____ lpm BVM without
intubation _____ BVM with intubation _____ Needle
decompression _____ Other _____

Response: No comment

The paramedic's management of this patient's breathing was: Optimal _____
Acceptable _____ Not Appropriate _____ Cannot tell

Response: No comment

Circulation
Was there a need to manage this patient’s blood pressure in the prehospital setting? Yes _____ No _____

Response: Think of circ more globally: volume, cardio status, etc.

Indicate reason for need to manage this patient’s blood pressure:
- Hypovolemia ______ Systolic BP < 90 ______ Signs of shock ______
- Burns to > 20% BSA ______
- Other Indication ____________________________________________

How was this patient’s blood pressure managed?
- IV Solutions ______ Number of IV Cannula ______ Amount of fluid ______
- Crew not IV endorsed ______

Response: No comment

The paramedic’s management of this patient’s blood pressure was
- Compliant _____ Not Compliant _____ with BCAS Hypovolemia Protocol.

Response: Referred to * above, and the idea behind the question made sense.

Cervical Spine

Was there a need to manage this patient’s cervical spine in the prehospital setting? Yes _____ No _____

Response: No comment

Indicate reason for need to manage patient’s cervical spine: LOC ______ Mechanism of Injury ______ Injury to Head/Neck

- Associated Injuries ______
- + Other Indications:
Response: Altered level of consciousness
Drugs and alcohol use with MOI independent of above

How was this patient's cervical spine managed? Cervical Collar ___ Backboard ___
Clamshell ___ Other ___

Response: No comment
The paramedic's management of this patient's cervical spine was:
Optimal ____ Acceptable ____ Not Appropriate ____
Cannot tell ____

Response: No comment

Other Orthopedic Injuries Evidence of injuries to patient's extremities Yes ___ No ___

Response: Straight forward

Was there a need to manage this patient's extremity injuries in the prehospital setting? Yes ___ No ___
Response: Straight forward

How was this patient’s extremity injuries managed?
Splint ____ Traction Splint ____ Reduction/Alignment ____

Response: Straight forward

The paramedic’s immobilization of this patient’s extremity injuries was
Optimal _____ Acceptable ____ Not Appropriate _____ Cannottell ____

Response: Straight forward

How was this patient’s pain managed:
Entonox ____ Morphine ____ Pain not managed ____

Implies that if you didn’t use pharmacological agent, you
didn’t manage the pain. What about immobilization? Pain not
managed pharmacologically.

The paramedic’s management of this patient’s extremity injuries was Compliant ____ Not
Compliant ____ with BCAS pain management protocol.
Response: okay

Code 99 The patient required diversion to a trauma facility based on MOI. Anatomical findings and Physiologic Presentation.

Response: clear

The paramedic Complied with Code 99 Protocol.

Response: clear
References


