TESTING THE RELIABILITY

OF A

NEWLY DEVELOPED SEDATION SCALE

BY

MICHAEL PETER NAMAKA, B.SC. PHARM

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Pharmacy
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

TESTING THE RELIABILITY OF A NEWLY DEVELOPED SEDATION SCALE

by

MICHAEL PETER NAMAKA

A Thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

ACKNOWLEDGEMENTS

To my dear wife Debbie Namaka, for all the endless love, confidence and support needed to successfully complete my Masters.

I would like to extend an utmost thanks to my preceptors Dr. Nicholas Honcharik, Pharm.D. and Dr. Ruby Grymonpre, Pharm.D. for their countless hours of assistance, that I required throughout this research project.

I would like to thank the core group members, AICU physicians and nurses, pharmacists and respiratory therapists, for all their time and ideas that played an integral part in the successful completion of this research project.

TABLE OF CONTENTS

ABSTRACT

A sedation tool for AICU (Adult Intensive Care Unit) patients is required to provide sound clinical care. Development of a new sedation scale was undertaken to overcome limitations of the available sedation scales. The goals of this research were to design a scale that possessed Content Validity with goal based parameters (with clearly defined parameter levels). In addition the tool was tested for reliability. To ensure the tool was reliable, the tool underwent interobserver reliability testing. An assessment tool composed of 9 goal based parameters was developed. A pilot study was conducted in 10 AICU patients to ensure that no gross problem areas were present. Following this, all minor alterations were completed. During the main trial phase, one parameter was assess in 211 AICU patients, while the remaining 8 parameters were assessed in 112 AICU patients.

The developed tool does possess Content Validity. The final results revealed that the scale as a whole "could" be deemed reliable based on the overall values for Percent Agreement (91.87%), Disagreement Rate (0.0813) and weighted Kappa (0.5137). Strigent reliability testing of each of the nine individual parameters revealed that 4 passed, 1 was borderline and 4 failed. Modification of the 5 substandard parameters is required in order to achieve complete reliability of the tool as a whole. Once this has been achieved, this tool, will have a significant impact on use of sedation in the critically ill patient.

PART 1: OVERVIEW OF SEDATION AND SEDATION SCALES:

Discomfort and Distress in Adult Intensive Care:

Postoperative patients in adult intensive care units (AICU) are under considerable discomfort and distress associated with their critical illness and injury. The highly variable and hectic environment of the AICU is also a contributing factor to distress. Examples of environmental stressors include: bright flashing lights, noise and activity associated with monitors and respirators, medical emergencies, lack of sleep, frequently changing caregivers, and a variety of painful procedures[1]. Discomfort and distress is further enhanced by the presence of indwelling catheters which require prolonged periods of immobilization. The stress of all this, compounded with the uncertainty as to their outcome, challenges even the most stable psyche.

Psychological distress is defined as a organism's response to aversive internal and external stimuli and may include discomfort, anxiety, fear and at the extreme, pain.[2,3] The organism's response is multidimensional and includes three components: a) behavioral (agitated movement, grimacing, crying etc.); b) physiologic (increased muscle tension, heart rate, blood pressure, hormonal response, etc.); and c) phenomenologic (self-report of anxiety, fear, pain,etc.). Managing discomfort or distress is an essential treatment goal and is given high priority by intensive care staff.[4] Inadequate clinical management of distress may interfere with the care of the patient and may also lead to detrimental consequences such as cardiac ischemia.[5,6]

Sedation:

Sedation, using various pharmacological agents, is a common measure taken to

minimize patient distress and/or discomfort. The word sedation is derived from the Latin word *sedare*, which means to make calm.[17] Sedatives, when given in the appropriate amounts, contribute to patient comfort, physiologic homeostasis, and safety.[7,29] Evidence indicates that use of sedative agents promotes sleep which in turn facilitates the healing process through restoration of damaged tissues.[7]

Over the last few years, the degree to which a patient should be sedated has been a controversial topic. In the past, deep sedation was the common standard practice. In a survey conducted by Merriman (1981), deep or complete sedation was considered as the ideal.[8] Sixty-Seven percent of the Intensive Care units aimed to keep most patients well sedated and detached from the intensive care environment. However, more recently it has been shown that heavy sedation is not a benign therapy and has several shortcomings.

Deep sedation cannot remove the cause of continuing stress. For example: a patient with septicaemia has a sixty percent increase in metabolic demands. Heavy sedation has been shown to reduce the metabolic state of these patients.[9] Other disadvantages of heavy sedation include: masking signs of neurological events, respiratory muscle fatigue, deep vein thrombosis, general muscle wasting, risk of compression injuries to peripheral nerves, infection, prolonged coma due to sedative overdose, and weaning problems.[10,11]

Deep sedation is now only reserved for a few specific indications. These include: patients with tetanus (where it may be used to suppress the autonomic effects), those with critical oxygenation (in whom a reduction of oxygen consumption may be

desirable) and patients with severe head injuries (where it is necessary to reduce cerebral oxygen consumption and to prevent coughing and straining which will increase intracranial pressure).[12]

In 1987, Bion and Ledingham conducted a survey to reassess the current perception of the ideal level of sedation.[13] Their results revealed that the ideas have drastically changed from the time of Merriman. Sixty-nine percent, regard the ideal level as asleep, but easily rousable. Other current perceptions describe optimal sedation as that point where a patient is sedated to the point of no distress with the use of medication on a as needed basis.[14] Recent reports have also defined the ideal level of sedation as the state where the patient is comfortable and spontaneously awakens from sleep or could be roused if required.[15,16] There is one common denominator linking the different definitions of the ideal level of sedation. This link is the concept of light sedation, leaving the practice of heavy sedation for the most part, an idea of the past.

However, there are risks associated with undersedation. Undersedating the critically ill may result in hypertension and tachycardia which may damage an already ischemic myocardium. It may also precipitate hypoxia, hypercarbia, barotrauma, and premature extubation.[18] As a result a balance between undersedation or oversedation must be reached. This brings forth the question: How to achieve and maintain an effective sedation level? One method is to utilize a tool to assess the level of sedation.

Sedation Scales:

In theory, the ideal sedation scale should be: valid, reliable, effective, quick

and easy to use, involve minimal workload and distraction to staff and present no additional hazard or discomfort to the patient. In addition it must also be sensitive to drug effects so that possible drug-induced changes in level of alertness can be detected, not require any apparatus, require little training of personnel, have well defined criteria and a standard method of administration, be based on clinically relevant criteria and not alter the level of consciousness of the subject during the assessment period. Also, of the utmost importance, it should define an optimal endpoint for titration of sedating drugs. This endpoint should correspond to adequate control of the clinical problems for which sedating drugs are given. This ideal scale must be suitable for recording at frequent intervals along with other routine indices of patient status. Over the years a variety of sedation scales have been developed for critically ill adult patients. These scales conformed to some of the above desired properties but none have ever been adequately been tested for reliability and validity.[19] To date, a scale composed of <u>all</u> these above properties does not exist for adult critically ill patients. However in pediatrics, the "Comfort Scale" **(TABLE 3A)** has been shown to be a valid and reliable assessment tool.[43]

Twenty years ago, Ramsay and colleagues developed a six point scale, to measure sedation by bedside observation.[20] This scale became the gold standard for assessing the level of sedation in the adult intensive care units. The Ramsay Sedation Scale, served as a template for the development of the various scales present today.[5,15,18,21-27] **Refer to TABLE 1** for some examples of scales used to date.

Problems of Existing Sedation Scales:

Although, existing sedation scales are being utilized in the adult intensive care setting, they possess several drawbacks which hinder their effectiveness. **See TABLE 2.** Before a sedation scale can be considered an effective clinical assessment tool, one must objectively answer the question: Does this instrument effectively measure the need for, or adequacy of, sedating drugs in an AICU patient? This question can not be objectively answered because none of the scales used in the AICU have been scientifically tested for reliability or validity. Reliability is defined as the extent to which a measuring procedure yields the same results on repeated trials. Validity has been defined as the extent to which a test measures what it is purported to measure. Reliability and validity are the essential components or cornerstones, that are required for the foundation of scale development. Without them, no matter what levels or parameters the sedation scale may possess, it becomes useless as a scientific measuring device. As a result, scales lacking these two critical components must be used cautiously, only as a rough guideline.

Furthermore in some scales, the various levels of sedation are not *mutually exclusive* from one another. This type of overlapping does not allow a complete, accurate assessment of sedation, and therefore hinders the scales full effectiveness.

Also, in some scales, the levels of sedation are *not clearly defined or fully inclusive*. This additional problem, enhances the already cloudy picture, of dependability and effectiveness. All scales *fail to define a clear endpoint* for sedation. This creates a major problem such that prospective comparisons between sedating

drugs can not be properly accomplished.[28,29]

Another problem lies within the very focus of the tool. Sedation scales are essentially a measure of consciousness, which are used for the purpose of achieving an "adequate sedation level", thereby preventing undersedation and oversedation. However, in the AICU, life saving measures sometimes require much deeper levels of sedation (example: severe head trauma). Also deep levels of sedation may be caused by some other process other than sedation. Hence, it becomes obvious that the "adequate sedation level" is not a universal phenomena. This brings forth the need to develop and establish a list of *exception criteria* (i.e. patients to be excluded) for scale use. Because each specific patient population may require a different level of sedation that is termed adequate, it stresses the importance of tailoring a scale to meet the specific demands of a specific patient population. At present, the existing scales have not taken these points into account and therefore should not be used universally in all patient population groups for all medical conditions.

Another major problem lies in the nature of the evaluating parameters present in the sedation scales. There are two types of measuring parameters which exist: Objective and Subjective. The difference between the two is that objective parameters are tangible while subjective parameters are not. Since there are no devices which monitor patient comfort and give a numerical readout, it is difficult to use physiological endpoints to monitor sedation. Hence, the monitoring parameters that comprise any sedation scale are primarily subjective. As a result, these scales rely heavily on the assessments of the clinician. At first glance, one might feel that

objective parameters would yield the most concrete results. In actual fact, objective parameters such as haemodynamic and respiratory variables in critically ill patients, may not be related to the level of sedation but rather to the disease process or drug effects.[18] Although the assessments conducted by clinical experts are primarily subjective, they should be examined from a objective point of view. This however does not eliminate *the possibility of observer bias*. Therefore methods used to reduce observer bias must be present within the scale design. The various sedation scales fail to take this in to account, thereby further diminishing the faith in the accuracy of assessing the patients level of sedation.

Coma recovery scales (i.e. Glasgow Coma Scale) are unsuitable for patients in the intensive care units as they use *inappropriate monitoring parameters*. They were designed as neurological assessment tools for patients with neurological deficits.[30] In AICU, immobilization of limbs by splints or pain as well as inability to verbally communicate due to the presence of a endotracheal tube, completely nullify the effective utilization of these scales. Although some parameters are very useful, other parameters used in these scales are simply not applicable to AICU patients and make complete assessment virtually impossible. For example, proper assessment of the eye opening component of the Glasgow Coma Scale in a patient with severe head injuries having swollen eyelids is virtually impossible.[33] Furthermore, in the Glasgow Coma Scale it was shown that error rates in assessment were highest at the intermediate levels of consciousness for which the detection of changes in the condition is vital.[32]

The present scales also fail to clearly define goals that they wish to achieve.

This lack of defined goals, prevents us from establishing content validity of the scale. If we don't know what its supposed to do, how are we to determine how effective/valid it is?

Some of these scales are also problematic, because certain monitoring parameters require the observer to *impose a specific arousing stimulus* to the patient in order to assess the degree of sedation. This type of arousing stimuli are intrusive and can influence the quality of sedation or presence of distress when frequently and repeatedly applied. Reflecting back to the "ideal sedation scale", it should present no additional hazard or discomfort to the patient nor should it alter the level of consciousness of the patient during the assessment period. It is quite clear that these two essential criteria are obviously not being fulfilled by some of the existing intrusive tools.

Lack of conciseness is also a problem in various scales that are too broad, inclusive, and cumbersome.[31] These types of scales often require intense training for their use. They are very *time consuming* and *difficult* to use and are therefore very impractical for the hectic environment of the AICU. In order to be practical and effective, the assessment via the scale should take less than 5 minutes to complete.

Some sedation scales present today, have one important common characteristic. They are all broken down into various components. In order for a scale to be statistically and scientifically sound, each component that comprises the scale must be *appropriately weighted* in terms of importance.[34] Otherwise, a favorable scoring in two or three areas of low importance can override a very poor scoring in another area

of high importance and hence inappropriately skew the final total sedation score.

Another possible way to avoid this weighting problem is to assess each of the various major components of the scale and assign the appropriate weight to each of the individual monitoring parameters that comprise that component of the scale. As a result, each component of the scale will be independently weighted from each other and henceforth have their own independent score. Independently assigning a score to each major component of the scale may yield more definitive information regarding that specific aspect of the patients status then an all inclusive, overall comprehensive score for all major scale components lumped together. Utilizing this technique, can immediately direct the patient caregivers to a specific problem area rather then just getting a generic number for the overall sedation score which indicates that something may be wrong with the patient, but yields no direction as to what it is.

Critical Evaluation of the Ramsay Sedation Scale:

To illustrate the various drawbacks and problems that were previously discussed, a detailed critical evaluation will be conducted on the **Ramsay Sedation Scale.** This scale was selected because it served as the template from which many other scales were developed. Despite it's widespread use, the Ramsay Sedation Scale is not an effective clinical assessment tool.[19] The Ramsay scale has never been scientifically tested for reliability or validity. This is a very significant drawback, because reliability and validity are two of the most fundamental aspects required in an effective sedation scale.

The Ramsay scale as shown in **TABLE 1A,** is composed of six levels of

sedation. These levels are not mutually exclusive from one another. For example, experienced medical staff are fully aware that mechanically ventilated patients can remain agitated with progressive sedation, even to the point of near unconsciousness.[19] As a result, a patient could be restless or agitated as shown in LEVEL 1 while at the same time be responsive only to a light glabellar tap or loud auditory stimulus as shown in LEVEL 4 or LEVEL 5. These six levels of sedation are not clearly defined or fully inclusive. For example, how can you distinguish the difference between a patient that is awake, cooperative, oriented, and tranquil as seen in LEVEL 2 from a patient that is awake, responding to commands only, as seen in LEVEL 3? Also, because each of these levels are not evenly spaced, average Ramsay scores are not meaningful.

Another major problem is the failure of the Ramsay scale to define any clear endpoint for sedation of patients. Upon examination of the scale, one might at first, select LEVEL 2 as the point of optimal sedation. This assumption couldn't be further from the truth. Some critically ill patients who are severely agitated or delirious cannot be brought to an awake, cooperative, tranquil state by the administration of the various available psychoactive drugs. These select patients require heavier doses of medication that place them in much deeper levels of sedation as illustrated in LEVELS 3,4,and 5. This clearly illustrates that an optimal level of sedation in one type of patient may not necessarily be the optimal sedation level in another. As a result, although LEVEL 2 may be an optimal sedation level for some patients, it is not the universal endpoint for all patients.

Besides exclusion criteria, a list of goals hoped to be achieved through adequate sedation is essential and has to be defined well before the scale is even developed. These goals are the fundamental components which are required in order to establish a clearly defined endpoint. The concept of "optimal sedation" can therefore be justified and validated according to whether these goals for a specific patient population are being met. It is important to note that these goals hoped to be achieved will change in accordance to the specific patient population under examination. As a result, optimal endpoints for sedation, will vary from one patient population to another. The Ramsay Scale does not take this into account.

The Ramsay Sedation Scale has been used by clinicians over the past twenty years. The problem with this type of assessment, is that it relies solely on the individual judgement and expertise of the person using it. Personal bias in this type of setting is unavoidable and therefore casts doubts on the value of this assessment tool. This shatters the scales foundation for dependability and reliability.

One more final issue needs to be addressed with respect to the Ramsay Sedation Scale. This problem lies in the fact, that this scale lacks the mandatory internal framework which is necessary to determine the sedation level of the patient. This scale presents with only the final product (i.e. different levels of sedation) without any workup as to the criteria that was assessed in order to place a patient in a given level.

The internal framework of an ideal scale should be composed of several major independent scale components. Each of these components should have several distinct

monitoring parameters that can adequately measure the goals that were initially established at the start of the scale development. Each monitoring parameter should be characterized by several specific and distinct weighted degradations of it from best to least. Each of the scales major components, should be independently assigned a score which should be determined by the weighted values assigned to the different graded levels of each of the monitoring parameters used to describe each scale component. Individual scale component scores should be obtained rather then an overall compounded sedation score derived from all scale components.

By applying the concepts above this will allow the caregiver to immediately detect a specific problem area in the patients status rather than just getting a generic number for the overall sedation score which indicates that something may be wrong but yields no direction as to what it is. Although a scale should be concise, it must be complete. The Ramsay Sedation Scale clearly does not meet these expectations.

The Ramsay Sedation Scale does however, present a few valuable aspects that should be noted. Implementing this scale does not require any appreciable invasive techniques that would impose any potential hazards or discomforts to the patient. The scale is very easy and quick to use, requiring minimal training.

The critical examination of the Ramsay Sedation Scale gives us hands on experience in recognizing the various problems the researcher must overcome. The Ramsay scale only passed three of the fourteen problem areas. This facilitates the need for the development of a new, valid, clinical assessment tool for sedation of AICU patients.

PART II: SEDATION SCALE DESIGN AND DEVELOPMENT:

Research Objectives:

a) Develop methodology to prove Content Validity of the scale and in the future explore Construct Validity.

b) To identify the behavioral and physiologic dimensions that are essential for rating sedation in AICU patients.

c) To develop an observer rating scale utilizing these dimensions.

d) Develop methodology to prove interobserver reliability of the scale.

Assisting the researcher was a *core group* of consultants consisting of: Dr. Nick Honcharik, Critical Care Clinical Pharmacist; Dr. Dean Bell, Anesthesiologist/Critical Care Physician; Dr. Colin Bands, Critical Care Fellow; JoAnne Major, Clinical Instructor; Critical Care Nursing Program; Louise Lemoine, Critical Care Nurse; Gail Hall, Critical Care Nurse; and Maurita Kiesman, Critical Care Nurse. In addition, Dr. Jennifer Clinch, WHO Quality of Life Section, St. Boniface Research Foundation and Dr. Thomas Hassard, Director of Biostatistics Unit, Professor of Biostatistics, Department of Community Health Sciences, University of Manitoba were also consulted periodically regarding scale development.

Content Validity:

The initial step in the assessment tool development was to obtain Content Validity. Content Validity has been defined as the extent to which a given test measures what it is purported to measure.[26] Tyler et al. described validity as a general approach to determine the extent to which a measurement tool conformed with

the clinical expectations.[58]

A thorough literature search was completed to compile a list of "Goals of Sedation" for AICU patients. In addition, specialty medical, nursing, pharmacy, and respiratory therapy staff were solicited in order to extract further crucial information. To ensure Content Validity of the scale, all the goals of sedation must be established and defined before actual scale development occurs. Various monitoring parameters were developed, to measure the attainment of the listed goals. The development of these monitoring parameters determined the internal framework of the developed scale.

In order to obtain the views of MICU/SICU staff, six different group discussions were conducted which encompassed virtually all MICU/SICU nursing, pharmacy and respiratory therapy staff. The researcher also scheduled a separate meeting with the critical care AICU physicians and fellows. As a result, a composite list of twenty-five goals of sedation was created (**TABLE 4**).

The next phase, involved the development of a method to extract the essential, goals of sedation, from the all encompassing list, derived from the various meetings. The researcher accomplished this through the use of a specific questionnaire (**TABLE 5**). Individuals ranked each goal in relation to the following statement: "clinically relevant and mandatory to ensure safe patient care". Ranking was documented as to the degree of priority (high, moderate, low) in the care of AICU patients. One hundred and sixty eight forms were distributed. To avoid bias, all participants filling out the form were informed to complete it independent of each other.

The response rates for physicians, nurses, respiratory therapists and pharmacists

were: 40%, 38.40%, 65%, and 87.50% respectively. The overall response rate for the survey was 45.83%. The results of the survey revealed 3 distinct separations of the goals listed in **TABLE 4**. Those goals which had the majority of respondents voting for low priority were immediately dismissed. However, those goals where the majority of respondents voted high priority or medium priority were kept and organized in decreasing order of importance **(TABLE 6)**. "High Priority" goals, ranked 1 to 10, received greater than 50% of the respondent votes. "Medium Priority" goals, ranked 11 to 14, were the next most important, capturing greater than 50% of the votes for their specific category. These goals were then separated into 3 categories that evolved due to commonalities between the goals. These categories included: Behavior, Comfort, and Therapeutics **(TABLE 7)**.

The list of goals were presented to the core group for further refining. After the initial refining process, goals #28, #15, & #25 were deleted. Goal #25 (*reduce aggression*) was deleted because it was incorporated into goals #1, 2 & 3. Goal#15 (*Optimize pulmonary hemodynamics*) was deleted, because this area was addressed in Goal#11 (*Optimize ventilation in patients who are asynchronous with the ventilator*). Finally, Goal #28 (*Reduce psychosis*), was deleted because it can be evaluated in part under Goal#3 (*Promote patient safety through reduction of destructive movement by patient*). Furthermore, a proper psychosis evaluation would require a separate scale. The purpose of this refining process, is to avoid any goal duplication.

Further refining resulted in all non-essential goals being deleted. Therefore, Goal #10, #12 & #13 (*Reduce intracranial pressure, Reduce myocardial oxygen*

demand and Optimize patients vital signs such as heart rate, blood pressure, respiratory rate) were deleted. All these goals are directly related to vital signs. Although they yield objective results, they are the poorest assessment tools to use as previously addressed. Furthermore, Goal #6 (*Prevent delirium tremors for patients undergoing alcohol withdrawal*) was deleted due to its extreme infrequency in the unit. Goal #24 (*Decrease level of consciousness in a paralyzed patient*) was deleted because paralyzed patients, were excluded, as per the exclusion criteria, that was established at the start of the study. **(See Reliability section)**. Goal #22 (*Comply with patients request for sedation*) was also deleted since it overlaps with Goals #18 & #5 listed under the Comfort Category.

A core group of essential goals remained which were separated into 4 distinct categories. The first category was developed solely from Goal #11 (*Optimize ventilation in patients who are asynchronous with the ventilator*). The next scale category/component to arise was Level of Consciousness. This component is not specific to any particular essential goal but is relevant to all essential goals that survived the sifting out process. The third scale component was Destructive Movement. This was developed based on Goals #1, 2, and 3. The final scale component, Comfort, was developed based on Goals #18, 5, 17, and 26.

In summary, the above described methodology represents the processes needed to ensure Content Validity. It is important to re-emphasize that in order to achieve Content Validity, the desired goals have to be defined upfront. Once this has been accomplished, specific monitoring parameters can be developed to properly assess

whether these goals are being achieved or not. This forms the key element in scale design and development.

Construct Validity, is also an important type of validity but it will not be addressed in this research project. Testing for Construct Validity (does the tool actually measure what its purports to measure) can take one or more years to complete and therefore will be the basis for another study.

Development of Behavioral and Physiological Monitoring Parameters:

From the initial template, [the 4 categories] the final tool was constructed. The parameters created for each of the 4 main categories previously discussed, evolved from an extensive literature search which identified all behavioral and physiologic parameters required to assess sedation in AICU patients. This search was then cross referenced with the concise list of "Goals of Sedation" mentioned in the section "Content Validity". Cross referencing these two searches, accounted for the addition or removal of certain monitoring parameters that will be present in the final draft of the sedation scale. All relevant monitoring parameters were placed under the respective category to which it belongs. Each monitoring parameter was further separated into distinct, mutually exclusive, degradations, from "worst" to "best" ensuring no overlap between each different degradation.

The monitoring parameters developed for the first scale component were: Oxygen Saturation, Degree of Asynchrony, Color, and Paradoxical Breathing (TABLE 8). These breathing parameters were specifically created to assess whether this goal is being attained or not. Specific definitions for "Asynchrony" and "Paradoxical

Breathing" were established to ensure uniform clarity during assessment. Each monitoring parameter was subject to extreme critique by the core group, to ensure that the various graded levels (from "worst" to "best") present, are unique and mutually exclusive. This would ensure no overlap between the graded levels.

The second scale component (Level of Consciousness) evolved into an actual parameter, composed of four, independent, levels. Each of the four levels, indicate the varying degrees of consciousness that are present **(TABLE 9)**. A specific definition was established for ("Painful Stimuli") to ensure uniform understanding of the term during assessment.

The third scale component (Destructive Movement) also evolved into a parameter. It was composed of four mutually exclusive, levels, which assess the various grades of Destructive Movement that are present **(TABLE 10)**. A specific definition for "Episode" was established to ensure uniform clarity during assessment.

For the final scale component (Comfort), two monitoring parameters evolved. These were: Degree of Facial Grimacing and Voice and Movement Gestures. These parameters were created in order for complete assessment of the different aspects of Comfort **(TABLE 11)**. Each of these monitoring parameters were composed of various graded levels from "worst" to "best" which were mutually exclusive, ensuring no overlap between the graded levels. A specific definition for "Oriented" was constructed to ensure uniform clarity during the assessment of "Voice and Movement Gestures".

The newly developed scale was then forwarded to the core group for further

refining. The researcher also solicited the bedside nurses via various on site interviews in order to capture their expert opinion on the newly developed tool. These interviews unveiled, the thought process of how each nurse actually assesses the different levels, of each relevant monitoring parameter present in the scale. Incorporation of this additional information resulted in the final version of the tool **(TABLE 13a, 13b, and 13c)**.

The assessment tool was devised such that a summation of scores from all parameters was not needed (i.e. a total score of all the parameters). Rather, the caregiver would only assess the parameter(s) which were related to the clinical problem present. This would strengthen goal related assessment by focusing on specific problem areas rather than conducting a overal global patient assessment. Furthermore, this type of focused assessment decreases the time required to use the tool.

Reliability Assessment:

Reliability is defined as the extent to which a measuring procedure yields the same result on repeated trials in the same study.[26] An extensive literature search was conducted to gain insight on how to determine scale reliability. This search included a detailed examination of sedation scales, coma scales[35,37], neurological scales[36], arousal scales[38], pain scales[39] and others.[40,41,42] Although there are various alertness/sedation scales developed for adults, there is only minimal information available on proper statistical methodology needed to prove scale reliability.[26] However, more extensive research in this area has been done in pediatrics.[43-51] **Refer to TABLE 3.** The

researcher did not incorporate the use of new statistics to assess reliability but did however reformulate new methods to utilize these statistics to their full potential.

The first phase involved the selection of the number and type of professional expert observers. One critical care physician and two experienced critical care nurses were required. Although most studies utilized only two expert observers, it was decided to have at least three, so that statistically more conclusive results could be obtained.

Following this, the researcher was involved in proper instruction and training to ensure proper utilization of the developed tool. Since this scale is concise, quick and easy to use, minimal training time was required.

Statistical Analysis for Reliability:

The first step towards sound, statistical analytic evaluation of reliability is to determine the minimum number of observations that need to be conducted. A biostatician was consulted to ensure that for all relevant statistics, proper analytical methods were being utilized to their maximum capabilities. Schefler has illustrated various statistical calculations which involve the number of observations required, correlation co-efficients, confidence limits and the degrees of freedom for this type of research under investigation.[61]

Although the developed sedation scale is a complete/whole functional unit, it is comprised of two main subdivisions that must be analyzed separately. The first division represents the bulk of the sedation scale and consists of all monitoring parameters *except* Destructive Movement. The second subdivision is Destructive

Movement. The rational for assessing Destructive Movement as an isolated Monitoring Parameter will be addressed in the section "Pilot Study".

For the first subdivision, the number of observations required by 3 independent, simultaneous observers was calculated using the Fisher Z Transformation technique. A minimum of 104 separate observation periods were required in order that this portion of the study be statistically sound. This was based on a reliability co-efficient of 0.8 with 95% confidence limits and 2 degrees of freedom.

Because the second subdivision, Destructive Movement, utilized only two observers instead of three, double the amount of observations were required in order for definitive results to be obtained. This value of 208 observations was calculated again through utilization of the Z Transformation technique. However it is based on a reliability co-efficient of 0.8 with 95% confidence limits with only 1 degree of freedom. The degrees of freedom were reduced due to the utilization of only 2 observers to assess this monitoring parameter. This shortcoming is compensated for in the formula by doubling the sample size of observations required.

The reliability of the scale is directly dependent on the degree of correlation of interobserver agreement that exists. Statistical methods that were used to assess interobserver agreement include: 1) Disagreement Rate, 2) Weighted k (kappa), 3) Percent Agreement. Both subdivisions will be evaluated using the above stated methods.

The Disagreement Rate as calculated by the Teasdale et al formula[52], results from discrepancies between individual ratings and consensus ratings (i.e. the modal

observations for that patient) is sensitive to the magnitudes of the disagreements between observers, not only to the number of disagreements. It indicates the average discrepancy between a single observation on a patient and the mode of the observations by all observers on that patient, expressed as a fraction of the maximum possible distance from the modal (consensus) rating. The main drawback with Disagreement Rate is that it tells you that a disagreement has occurred, but it does not tell you how far apart or the extent to which that disagreement occurs. For example: with respect to Destructive Movement, if the primary nurse selected level 1 but the secondary nurse selected level 5, it becomes clear that a disagreement has occurred but it is evident that this statistic does not capture the magnitude of this disagreement.

In weighted k[53], the weights are assigned according to those described by Fleiss.[54] A "k" value of 0 reflects only chance agreement. Positive values for "k" indicate greater than change agreement.[55] A generally accepted guideline for interpreting "k" has been proposed by Landis and Koch.[56] They propose that "k" less than 0 indicates poor agreement, 0 to 0.2 indicates slight agreement, 0.21 to 0.40 indicates fair agreement, 0.41 to 0.60 indicates moderate agreement, 0.61 to 0.80 indicates substantial agreement and 0.81 to 1.00 indicates almost perfect agreement. **Refer to TABLE # 14** Weighted K (Kappa) SUMMARY. One very important limitation of Kappa is that Kappa is, in fact, an index of "shared discrimination", rather than an index of agreement. In other words, it measures the ability of raters to consistently differentiate between diagnostic situations.[62] In situations where all raters agree that only one diagnosis occurs in the entire sample, Percent Agreement will be

100%, but Kappa cannot be calculated because no discrimination occurs in the study. In a situation where one single diagnosis occurs overwhelmingly, the Percent Agreement will be very high but Kappa will usually be very small, since there is relatively little natural discrimination in the data set. In other words, in order for the Kappa statistic to properly assess reliability, there has to be a reasonable spread/distribution of the data collected with respect to the percentage of time, that a given level was selected. Otherwise, if the data is skewed dominantly to one side, the Kappa statistic becomes a useless method to assess reliability.

Percent Agreement is a widely understood method that is also used. It is an all or none measure and is not sensitive to the amount of variability among the patients being observed or to the magnitudes of the errors made.[32] Greater than eighty percent agreement indicates relatively consistent agreement.

In summary, if the Disagreement Rate is less than 0.22, the Weighted "k" is greater than 0.4 and the Percent Agreement is greater than or equal to eighty percent, then it can be stated that there is a high correlation of interobserver agreement. As a result, the researcher could conclude that the scale is a reliable clinical assessment tool. If statistical analysis of the scale fails to conform to all of the above measures then it can be stated that there is a low correlation of interobserver agreement and the researcher could conclude that the scale is not a reliable clinical assessment tool. However, if statistical analysis of the scale, fails to conform to some of the above measures, then the researcher cannot definitively conclude that the scale is reliable and therefore must reassess the various monitoring parameters of the scale.

Pilot Study

A pilot study was undertaken to determine if any major problems existed (i.e. parameter gradations were clear and no overlapping was present). The sedation tool was assessed via a 10 patient pilot study by 3 members of the core group. The 3 observers were trained independently, over the same time period. Independent observation by each observer was essential to avoid any interobserver bias. It is important to note, that throughout this study, pharmacologic sedation was not administered by the observers, investigators, or bedside nurse for the purposes of this study. If a medication with sedative properties was administered prior to the observation session, the subject was considered a candidate for the study unless exclusion criteria were present.

The 10 patients assessed in the pilot were arbitrarily selected. Different patients were selected to test the outer limits of the scale. Variation or inconsistency in the scale is more likely to be seen by assessing different patients versus using the same patient at a different point in time. For the pilot, the same 3 observers were used for all 10 observations, therefore maintaining an internal consistent environment for scale use. As a result, any variation between observers can be directed more to the scale rather then to the observers.

The subjects were adult ICU patients within the Medical and Surgical ICU's of the Health Sciences Centre. Exclusion criteria include: patients on neuromuscular junction blocking agents to induce paralysis, paralyzed patients, jaundiced patients, liver failure patients, patients with seriously compromised neurological status, (severe

head injury, ischemic encephalopathy, stroke, status epilepticus), patients with altered muscle tone and those patients with critical oxygenation. In this way, we are clearly defining that this sedation scale is not a universal assessment tool applicable for all patient types.

This pilot study was unique, from the perspective that the 3 observers not only independently selected the appropriate level for each monitoring parameter but also commented whether 1, greater than 1 or no levels were applicable to the patient under observation. **See TABLE 12a & 12b.** The purpose of this additional assessment step was to ensure that the developed scale had mutually exclusive levels for each of the monitoring parameters present in the scale. In this way, if the observer found that more than 1 level or no levels were appropriate for the patient, then a problem existed and that aspect of the scale would have to be reevaluated. To address possible problem areas, the original form of the pilot study included an area for "discussion". In this way, the researcher could utilize comments to rapidly correct specific problem areas of the scale. This additional workload during assessments resulted in the longer (10 minute) assessment period.

Pilot Study: Results and Revisions:

The pilot study was a complete success, revealing no major problems. There was an excellent correlation for overall interobserver agreement at 86%. Only minor revisions were required and will be discussed below.

From the pilot study, it became evident that it is imperative to have all relevant working definitions placed directly on the bottom of the final draft of the sedation

scale. This change allowed the definitions to be readily accessible to the observer thereby avoiding confusion which could lead to improper scale use.

Furthermore, Destructive Movement was removed from the main sedation scale and assessed as a separate, entity. Originally, this scale component was to be assessed by the patient's bedside nurse. However, to avoid individual bias, this was not possible. In order to overcome this drawback, the core group constructed a plan to remove this type of bias. In the intensive care unit, patients are separated from each other only by a curtain partition. As a result the two nurses caring for each of their respective patients are in constant exposure to the status and well being of both patients. Therefore, the core group decided that besides the primary caregiver, the secondary caregiver should also be consulted to evaluate the Destructive Movement status of the patient. It is imperative to emphasize that the secondary caregiver, was the nurse that worked next to the primary caregiver, on a separate patient of their own, but relieved the primary caregiver for breaks or helped the primary caregiver's patient at different time intervals throughout the day. Therefore, comparisons can be made as to the degree of agreement that exists with respect to Destructive Movement between the primary caregiver and the secondary caregiver via the Kappa statistic. According to the Fisher Z transformation technique, it was determined that at least 208 observations would be required to yield a statistically sound result for Kappa. This calculated value of 208 observations is based on a reliability co-efficient of 0.8 with 95% confidence limits and 1 degree of freedom.[61] This new change in study design will eliminate the criticism of this potentially problematic type of observer bias.

Another important change from the original design was to have all 3 observers assess Degree of Asynchrony, Paradoxical Breathing, and Accessory Muscle Use at the same time over a 5 minute period. It was decided to assess these 3 areas over the same 5 minute interval because they are all functionally related to each other.

Furthermore, a specific order of assessment was developed. This was to ensure that all observers were assessing the same monitoring parameter at the same point in time. The core group decided that the Oxygen Saturation would be first, followed by [Degree of Asynchrony, Paradoxical Breathing and Accessory Muscle Use], Patient's Color (3rd), Level of Consciousness (4th), Voice and Movement Gestures (5th), Degree of Facial Grimacing (6th), and Destructive Movement (7th) as per primary RN and secondary RN respective forms.

Another important change that arouse was the bedside nurses involvement, in the assessment of the patients Level of Consciousness. Originally, in the pilot study, each observer took their turn to administer a painful stimuli to the patient or used their own method of trying to rouse the patient. In order to standardize these 2 techniques, all bedside nurses were trained to deliver a painful stimuli, when needed, in accordance to the definition and criteria established by the researcher. By having the bedside nurse administer the painful stimulus, the observers were all free to assess the patient's response at that specific point in time. Furthermore, it was decided to utilize the primary caregiver to try and rouse the patient because the patient is more likely to respond and be cooperative to familiar touch and voice of the primary caregiver who has already developed a rapport with the patient. Standardized methods for patient

rousal were developed by the core group and the bedside nurse was appropriately educated. The standardized method for patient rousal were as follows: the patient rouses spontaneously, if this is not the case the primary caregiver will then try to rouse the patient via soft voice and light touch; if this is still unsuccessful, the primary caregiver will then try to rouse patient via loud voice and light shaking of arm, if rousal still is not achieved then the primary caregiver uses a loud voice with painful stimuli. By establishing a consistent method of rousal for all patients, one can maximize the full effectiveness of the scale.

With respect to Level of Consciousness, in order to standardize the assessment of whether a patient can obey commands, all the bedside nurses were trained to ask the patient to "open your mouth" and "stick out your tongue", provided the patient was oriented and coherent. This newly adapted change led to more accurate and standardized assessment of Level of Consciousness.

From the pilot, it was also decided that the primary caregiver would be responsible for turning the patient (if no contraindications were present) to allow the observers free access to assess Degree of Facial Grimacing (i.e. discomfort). The primary caregiver was also selected to be responsible for initiating the patients self-assessment of discomfort via the visual analogue scale. All nursing staff in the Medical and Surgical ICU's were trained to familiarize them to the study and the assessment tool.

After the above revisions, the tool was ready for reliability testing, which was the basis for this investigation.

The Final Implementation Phase:

In this phase, the revised scale underwent reliability testing. The testing that took place was two dimensional. The first dimension examined and tested was the main bulk of the sedation scale, which consisted of all monitoring parameters *except* Destructive Movement. The second dimension examined and tested was Destructive Movement, as an isolated entity. The rational for assessing Destructive Movement as an isolated Monitoring Parameter was addressed under section "Pilot Study".

To assess observer agreement (reliability) in the first dimension, each subject was independently assessed by 3 observers: one physician and 2 experienced critical care nurses. Observers were obtained from a pool of 4 critical care physicians and 6 critical care nurses. Each individual was trained in the use of the sedation assessment tool, and the methods and procedures of the study. A day of the week was picked randomly where all 3 observers would be present. Subsequent observation sessions occurred 1 to 2 times a week (minimum of 3 days between observation periods) until all 112 observations were obtained. Patients were excluded from this study based on the exclusion criteria established and described earlier in the pilot study section. An interpreter was utilized for non-English speaking patients.

One very important caviat worth mentioning, is that multiple observations in the same patient was allowable because the time interval between observations was long enough (minimum 3 days) to ensure that duplicate observations are not obtained. Furthermore this type of data collection was possible because the ICU patient is often in a dynamic state.

The bedside nurse was instructed to perform proper rousal techniques, proper standardized utilization of the Visual Analog Discomfort Scale for Voice and Movement Gestures, and appropriate patient manipulation for proper assessment of Facial Grimacing as outlined under the section entitled "Pilot Study". It took roughly 3 minutes to instruct each bedside nurse in order to have complete standardized assessment of the patient by the 3 independent observers. During the actual assessment of patient, the researcher prompted the trained bedside nurse to perform the required duties needed in order for complete accurate assessment of the patient by the 3 independent observers.

The reliability testing of the second dimension (Destructive Movement), required a different study design to gather the appropriate data. The researcher utilized a 2 form technique, whereby one form was given to the primary caregiver and a second identical form was given to the secondary caregiver. They were both asked to select which level of Destructive Movement the patient is best described by. The forms filled out by both the primary and secondary caregivers were completed independent of each other in order to avoid observer bias.

Reliability Results:

The reliability assessment for the first scale dimension involved 112 observations. The 112 observations were captured over a 35 day interval involving 9 actual days of observation. Each of the 112 observations took roughly 8 minutes on average for complete patient assessment. The reliability assessment of the second scale dimension "Destructive Movement", involved 208 observations. Of the 320 total

observations required, only 3 observations involved the use of an interpreter (aboriginal).

The final results revealed an overall Percent Agreement of 91.87%, Disagreement Rate of 0.0813 and a Kappa of 0.5137. Stringent reliability testing of the nine individual parameters, revealed that 4 passed, 1 was borderline and 4 failed. The 4 passing parameters were: Oxygen Saturation, Level of Consciousness, Degree of Facial Grimacing, and Voice and Movement Gestures. Percent Agreement, Disagreement Rate and Weighted Kappa values for these parameters are illustrated in **TABLE 15**. These 4 parameters can conclusively be deemed reliable, passing on all 3 statistical methods. Destructive Movement was at best, a borderline pass for reliability, yielding a substandard Percent Agreement of 68.72%. The Disagreement Rate (0.3128) for this parameter was much higher than the standard acceptable value of 0.22. Weighted Kappa was also slightly substandard at 0.3889. The 4 failing parameters were: Degree of Asynchrony, Color, Paradoxical Breathing, and Accessory Muscle Use. Although, Degree of Asynchrony yielded a passing grade for Percent Agreement and Disagreement Rate, it fell well below the standard acceptable Kappa value of 0.4 by yielding a value of 0.3062. Color, Paradoxical Breathing and Accessory Muscle Use were analogous to that of Degree of Asychrony. These parameters showed excellent Percent Agreement and Disagreement Values. Their downfall, which led to their ultimate failure, lies in their substandard Kappa values. **See TABLE 15.**

Discussion:

The statistical analysis utilized for this study was unique from all other sedation scales. This uniqueness, arouse not in the type of statistics used, but the fact that these statistics were calculated and analyzed separately, for each individual monitoring parameter of the scale. In essence, each individual parameter must pass all 3 analytical methods in order to be deemed reliable. This type of accountability for each individual monitoring parameter, represents a very stringent method to ensure complete scale reliability. This is in contrast to reliability testing by other scale designers, where reliability was evaluated according to an average value (obtained for each statistic calculated), for all monitoring parameters present in the scale **(see TABLE 3)**. According to this rational as long as this "average value" for each calculated statistic, passes, then the scale is said to be reliable. This is quite misleading, because individual analysis of each separate monitoring parameter, may uncover parameters which have actually failed. As a result, this type of "averaging" leaves the reader of the literature, with a false sense of security with respect to the entire reliability of the scale.

Furthermore, other reliability testing conducted by some researchers, only analyzed the data via Percent Agreement and/or Disagreement Rate or Weighted Kappa. Study designs utilizing this approach are incomplete and yield only inconclusive results. Analysis of this sort, makes it easier for poorly constructed scales to be granted reliability status undiservingly. Proper, complete analysis, can only be achieved if all 3 statistics are used together on the collected data.

For example, in this study, the overall average for Percent Agreement, Disagreement Rate, and Weighted Kappa values were: 91.87%, 0.0813 and 0.5137 respectively, These "average values" are well above the standards established for reliability. However, one cannot automatically conclude the scale as a whole is reliable. Close examination of each individual parameter, revealed that only 4 out of the 9 actually passed, leaving 1 parameter as borderline, while failing the remaining 4 parameters. **See TABLE 15.**

The Monitoring Parameters Oxygen Saturation, Level of Consciousness, Degree of Facial Grimacing, and Voice and Movement Gestures successfully passed all 3 statistical analytical methods and can be deemed reliable. Destructive Movement was at best borderline while the Breathing Parameters, Degree of Asynchrony, Color, Paradoxical Breathing and Accessory Muscle Use, all failed. A detailed discussion examining each of these parameters will be addressed below.

Voice and Movement Gestures:

Voice and Movement Gestures yielded the highest overall value for Percent Agreement and Weighted Kappa (99.71% and 0.9873 respectively) while displaying the lowest Disagreement Rate (0.0029). This monitoring parameter illustrated an excellent spread/distribution between the two levels available for selection by the observers. Level 1 was selected in 37.80% while Level 2 was selected in 62.20% of all observations. This spread between the two levels is a critical factor (as discussed under Statistical Analysis) which accounts for its excellent Kappa value. The high Percent Agreement Rate with corresponding, extremely low, Disagreement Rate were

exactly what the researcher had expected. These expected findings were attributed to the standardized assessment techniques developed, which made it easier for the observers to determine whether or not a patient is alert, oriented and coherent as per Level 1. Once Level 1 is established, we can proceed to Level 2, to illicit the patients degree of discomfort, via the Visual Analog Scale. The numerical value obtained is based on the patients verbal or non-verbal response.

Oxygen Saturation:

Oxygen Saturation was the second most reliable parameter. Although an excellent Kappa value was obtained (0.9553), it was slightly lower then the previous parameter. The reason for this, can be illustrated again, upon examining the spread between the 4 levels that were available for selection. The percentage of selection, for each level were as follows: Level 1 (0.0%), Level 2 (2.08%), Level 3 (24.40%) and Level 4 (73.51%). These results clearly demonstrate, a less uniform spread, between levels, for this data set. As stated earlier, it is this spread of the data that directly impacts the calculated Kappa value for this parameter.

Percent Agreement was high (99.12%) with a corresponding low Disagreement Rate (0.0088). These results were expected due to the fact, that the observers obtained a numerical value from the Oxygen Saturation Monitor. The Percent Agreement and Disagreement values were slightly lower and higher respectively, than that of Voice and Movement Gestures, because the observers may have seen a different value for the Oxygen Saturation as indicated on the monitor. Poor timing of the exact monitor reading by the 3 independent observers, could account for this discrepancy.

Degree of Facial Grimacing:

The third most reliable monitoring parameter was Degree of Facial Grimacing. Although the calculated Kappa value was considerably lower (0.6858), than that of the previous 2 parameters, it is still, well above, the standard accepted Kappa Value of 0.4, used for reliability testing. The percentage of selection, for level 1, 2, and 3 were: 5.06%, 22.92%, and 72.02%, respectively. Although there was reasonable spread between the 3 levels available for selection, it was not sufficient enough to illicit the high Kappa Values obtained for Voice and Movement Gestures and Oxygen Saturation.

Percent Agreement (92.92%) and Disagreement Rate (0.0708) were lower and higher, respectively, than the above monitoring parameters, but still yielded excellent statistical values, well above the minimum acceptable standard of 80% for Percent Agreement and less than 0.22 for Disagreement Rate. One possible reason why lower results were obtained for these two statistics, lies in the numerous distractions in the AICU environment. As a result, simultaneous observation of a patient's facial grimacing, during manipulation of the patient by bedside nurse, did not always, consistently occur. Lack of focus by any 1 of the 3 observers, could have resulted in the higher discrepancies that occurred between the observers for this parameter. Another possibility is that the various levels for this parameter are not clear enough. As a result, the observers may not have fully understood the levels.

Level of Consciousness:

Although, Level of Consciousness fell behind, in comparison to the previous 3

parameters, it received a passing grade for reliability. The shrinking Kappa Value of 0.6719 can once again, be attributed to the lack of uniform distribution of selection between the 4 available levels. The results clearly indicate a more evident skewing of the distribution toward Level 4 at 70.83%. The respective percentages for Level 1, 2, and 3 are as follows: 7.44%, 13.99% and 7.74%. Although the Kappa Value obtained is well above the acceptable standard, but could have been higher if more uniform distribution of selection between the 4 levels was attained via assessing more variable types of patients.

The lower but still excellent value for Percent Agreement at 90.54%, can be attributed to the adequate standardized arousal techniques, taught to the bedside nurses by the researcher. This approach, supplied the 3 independent observers the consistency needed to properly assess which Level of Consciousness the patient represents. This same explanation accounts for the higher but still adequate value for Disagreement Rate (0.0946). This monitoring parameter yielded lower and higher Percent Agreement and Disagreement Rates respectively, than that of the other 3 monitoring parameters described above. This may have resulted from the complexity and subjective nature of this parameter compared to the rest of the parameters that comprise the scale. As a result, the researcher is pleased that a passing grade for reliability can be given to such a highly complex parameter.

Destructive Movement:

Statistical analysis of the parameter Destructive Movement resulted in a failing grade with respect to reliability. The results obtained for all 3 analytical methods were

substandard. The data revealed insufficient spread of selection between the 5 levels available, which resulted in the substandard borderline Kappa Value of 0.3889. The results clearly illustrate the skewing of selection toward Level 5 at 50.47% leaving Level 1 (10.66%), Level 2 (6.87%), Level 3 (10.90%), and Level 4 (21.09%) with sparse selection percentages.

As a result, minor alterations are necessary, in order to achieve a larger Kappa statistic, thereby strengthening the reliability status of this parameter. One minor adjustment, that could enhance the reliability results for this parameter, is the actual timing for assessment. Throughout the course of this study, the researcher presented to the AICU at around 11:00 a.m. and approached the primary and secondary nurse independently, to complete the patient assessment forms for Destructive Movement. The problem encountered was that, although the primary nurse had received report on the patient from the night nurse, the fact is, she only actually cared for the patient for a total of only 4 hours. A second major drawback is the fact that the secondary nurse has only cared for the patient during the 20 minute morning break and as a result a full proper assessment of the patient was still not attained. The initial assessment of Destructive Movement was to be conducted over the course of a 12 hour shift. Changing the assessment time towards the end of the 12 hour shift when both nurses can give a more complete and accurate patient assessment may improve reliability.

This same modification to the study design for data collection, could also possibly correct the substandard low and high values for Percent Agreement and Disagreement Rate respectively. The Percent Agreement Rate bottomed out at 68.72%

in conjunction with a corresponding high Disagreement Rate which peaked at 0.3128. In comparison to all other monitoring parameters, the results obtained for these two statistics were the worst overall. Percent Agreement and consequently Disagreement Rate values are much more difficult to obtain because of the number of observers utilized to assess this parameter. For this parameter, 2 observers were utilized versus 3 observers for all other monitoring parameters. This fact becomes a major factor upon examining the calculations that make up this statistic. When utilizing only 2 observers, Percent Agreement becomes an all or none value versus using 3 observers. For example, if 2 observers agree on a single assessment, 100% agreement exists. However if the 2 observers disagree, then it can be stated that 0% agreement occurs. In this case, there is no Percent Agreement between these extreme values. However, if 3 observers are utilized and 2 agree and 1 disagrees, it becomes quite clear that a middle range buffer, for Percent Agreement exists at 67% eliminating the all or none phenomena as described for the 2 observers scenario. This fact alone, makes it much more difficult to achieve high Percentage Agreement. This same analogy can be stated with respect to Disagreement Rate.

The Failed Monitoring Parameters:

All 4 breathing parameters: Degree of Asynchrony, Paradoxical Breathing, Accessory Muscle Use, and Color failed the reliability testing. This failure, results from the fact that a passing grade was not obtained for all 3 statistics utilized in this study. It does not mean that these parameters failed all 3 statistics. In actual fact, all breathing parameters surpassed the minimum requirements for Percent Agreement and

Disagreement Rates, but failed on the all important Kappa statistic. **See TABLE 15.** As outlined in our initial plan for data analysis, a parameter would not be deemed reliable unless it passes all 3 statistical measures. This type of stringent analysis makes it very difficult for a monitoring parameter to be granted reliability status.

After examining all the data, the researcher identified a common link that accounts for their failure. This link or commonality, lies in the distribution or spread of the data collected in the sample. As stated earlier, in order for the Kappa statistic to properly assess reliability, there has to be a reasonable spread/distribution of the data collected with respect to the percentage of time, that a given level was selected. Otherwise, if the data is skewed dominantly to one side, the Kappa statistic becomes a useless method to assess reliability. Upon examining this distribution of the data, it is quite evident that all data in the above stated parameters are tremendously skewed to one side. **See (TABLE 15).** For example, for Degree of Asychrony, Level #4 was selected in 95.36% of all observations compared to 0.61%, 0.0%, and 3.87% for Level #1, 2, and 3 respectively. This similar analogy can also be demonstrated for Color, Paradoxical Breathing and Accessory Muscle Use. For skewed data, the Kappa statistic is driven down. In all these cases, although the Percent Agreement and Disagreement rate are excellent and well above and below respectively the standard acceptable values, the Kappa statistic is the major player which determines whether a parameter is reliable or not.

As a result of this major drawback, these 4 parameters can not be considered as reliable. These parameters require major revision before they can be brought forth for

reliability testing again.

To overcome this grey area of the sedation scale, the core group was consulted to explore new ideas and suggestions for future studies. Firstly, color adds no value to assessment of patients breathing status and should be deleted.

Secondly, it has been stated by nursing and medical staff, that Paradoxical Breathing and Accessory Muscle Use are an all or none phenomena. As a result, a possible suggestion, is to eliminate the various levels present for each of these two monitoring parameters and simply convert this assessment to either a selection of "yes" it is present in the patient or "no" it is not present in the patient. If this technique is utilized, these parts of the sedation scale would have to be retested for reliability before actual implementation in AICU could take place. This suggestion may or may not mend the reliability problem. New more radical approaches may need to be developed in order to achieve appropriate assessment of these important breathing parameters.

Thirdly, with respect to Degree of Asynchrony, this whole monitoring parameter needs to be revisited, retested and modified in order to develop a truly reliable monitoring parameter. In fact, it may need to be incorporated into Accessory Muscle Use and/or Paradoxical Breathing rather than stand alone, as a solo monitoring parameter.

Scoring System:

Since an important part of this research was to test the reliability of a newly developed sedation scale, the actual scoring system for the scale utilizing assigned

weighting will not be conducted at this time. At present, because the different gradations for each parameter are organized from "worst to best", an actual scoring system *may not* be even required. However, proper methodology for developing a potential scoring system for this scale is addressed under the section "Critical Evaluation of the Ramsay Sedation Scale".

Conclusion:

Based on the overview of the literature on sedation and sedation scales, several very important conclusions can be drawn. Firstly, appropriate sedation of AICU patients enhances the healing process, promotes patient safety and helps to eliminate discomfort/distress factors that may further complicate the patient's general state of health. Sedation scales are the key to more effective utilization of sedation. It is hoped that once proven reliable, they will dramatically influence which drug will be used, the amount of drug used, the frequency at which it is given, and the duration of its use. From a thorough review of the literature, several problem areas with the various present sedation scales arose, as illustrated earlier, in the section entitled, *Critical Evaluation of the Ramsay Sedation Scale*. Methods were devised to overcome them and a new sedation scale with goal based parameters was developed. This tool was tested on AICU patients and statistically analyzed for reliability and validity.

The final analysis of this newly constructed tool, revealed several fundamentally important conclusions, with respect to validity and reliability. Based on the methods utilized for validity, one can conclude that this sedation scale does possess *Content Validity* but not Construct Validity. Testing for Construct Validity (does the

tool actually measure what it purports to measure) can take one or more years to complete, and therefore will be the basis for another study.

The results of this study reveal that the monitoring parameters: Oxygen Saturation, Level of Consciousness, Degree of Facial Grimacing, and Voice and Movement Gestures are reliable assessment parameters. However, the monitoring parameter Destructive Movement, is at best, borderline to being accepted as reliable. With some minor adjustments, this parameter could also be deemed reliable. However although this is encouraging, once modifications are made, future retesting must take place again, according to the outlined methods and analysis previously described, to ensure that it will be a statistically reliable monitoring parameter.

All Breathing Status Parameters: Degree of Asynchrony, Paradoxical Muscle Use, Accessory Muscle Use and Color, were shown not to be reliable and can not be depended on for conclusive results with respect to the breathing status of the patient. Further refinement of these parameters is necessary for the sedation scale to be a complete reliable entity.

Lastly, one can not state a scale is reliable based on the overall results. The overall results, although favourable, are quite deceiving. Only, after a stringent individual analysis, of each individual monitoring parameter can the true weak links in the scale be revealed. The parameters failing must be reassessed and should not be blended in and averaged with other parameters, in order to falsely claim the scale is reliable.

One can also conclude that, although the methods used to construct this

sedation assessment tool, addressed the problems depicted in the literature, it also uncovered new problem areas.

It is the researcher's goal to eventually have this clinical assessment tool incorporated as part of all routine assessments and/or charting done in these types of patients. Although more refinement of this tool is needed, the information obtained and shared during this research project will bring us one step closer to the elusive perfect sedation scale.

TABLE 1 (A - D).

A) The Original Ramsay Sedation Scale per *RAMSAY et al.*[20]

Level 1	Anxious and agitated, or restless , or both
Level 2	Cooperative, oriented and tranquil
Level 3	Responding to commands only
Level 4	Brisk response to light glabellar tap
Level 5	No response to light glabellar tap

B) The Present 7 Point Cambridge Sedation Score = Addenbrooke's Hospital Intensive Care Unit Sedation Scale per *O' SULLIVAN et al.*[18]

Agitated			
Awake			
Roused by Voice			
Roused by Tracheal Suction			
Unrousable			
Paralysed			
Asleep			

C) Modified Glasgow Coma Scale per *COOK et al.*[21]

	SEDATION SCALE
EYES OPEN	
Spontaneously	4
To Speech	3
To Pain	2
None	1
RESPONSE TO NURSING PROCEDURES	
Obey Commands	5
Purposeful Movement	4
Non-purposeful Flexion	3
Non-purposeful Extension	2
None	1
COUGH	
Spontaneous Strong	4
Spontaneous Weak	2
On Suction Only	2
None	
RESPIRATION	
Obey Commands	5
Spontaneous Intubated	4
Simv/Triggering	3
Respiration Against Ventilator	2
No Respiratory Efforts	1

D) Alertness/Sedation Scale per *CHERNIK et al.*[26]

1	**Responsiveness:** Assess responsiveness by calling the subjects name once or twice in a normal ton of voice. If the subject does not respond, call more loudly and/or repeatedly. If he/she still does not respond, mildly prod or shake the subject.
2	**Speech:** Ask the subject to repeat the sentence "The quick brown fox jumps over the lazy dog". Assess his/her speech
3	**Facial Expression:** Assess degree of facial relaxation
4	**Eyes :** Evaluate the subject's ability to focus and ptosis

TABLE 2

Summary of problems of existing scales :

#1	Lack of reliability
#2	Lack of validity
#3	Levels of sedation are not mutually exclusive from each other
#4	Levels of sedation are not clearly defined or fully inclusive
#5	Lack of a clear endpoint
#6	Lack of development of exception criteria
#7	Possibility of observer bias
#8	Inappropriate monitoring parameters
#9	Failure to clearly define goals of sedation
#10	Presence of invasive monitoring parameters
#11	Lack of conciseness
#12	Difficult to use
#13	Time consuming
#14	Lack of appropriate weighting of different levels of the scale

Table 3 (A - I) The Comfort Scale per *AMBUEL et al.*[43]

ALERTNESS	SCORE	TIME	*BLOOD PRESSURE*	SCORE	TIME
Deeply Asleep	1		Blood Pressure below baseline	1	
Lightly Asleep	2		Blood Pressure consistently at baseline	2	
Drowsy	3		Infrequent elevations of 15% or more during observation period	3	
Fully Awake	4		Frequently elevations of 15% or more above baseline(more than 3 during observation period)	4	
Hyperalert	5		Sustained elevations of 15% or more	5	
CALMNESS/AGITATION	SCORE	TIME	*RESPIRATORY RESPONSE*	SCORE	TIME
Calm	1		No coughing and no spontaneous respiration	1	
Slightly Anxious	2		Spontaneous respiration with little or no response to ventilation	2	
Anxious	3		Occasional cough or resistance to ventilator	3	
Very Anxious	4		Actively breathes against ventilator or coughs regularly	4	
Panicky	5		Fights ventilator, coughing or choking	5	
PHYSICAL MOVEMENT	SCORE	TIME	*HEART RATE*	SCORE	TIME
No Movement	1		Heart rate below baseline	1	
Occasional Slight Movement	2		Heart rate consistently at baseline	2	
Frequent , Slight Movements	3		Infrequent elevations of 15 % or more above baseline(1-3 during observation period)	3	
Vigorous Movement Limited to Extremities	4		Frequent elevations of 15% or more above baseline (more than 3 during observation period)	4	
Vigorous Movements Including Torso and Head	5		Sustained elevations of 15% or more	5	

Table 3A) **The Comfort Scale** per *AMBUEL et al.* continued

MUSCLE TONE	SCORE	TIME	*FACIAL TENSION*	SCORE	TIME
Muscle Totally Relaxed No Muscle Tone	1		Facial muscles totally relaxed	1	
Reduced Muscle Tone	2		Facial muscle tone normal. No facial muscle tension evident.	2	
Normal Muscle Tone	3		Tension evident in some facial muscles	3	
Increased Muscle Tone and Flexion of Fingers and Toes	4		Tension evident throughout facial muscles	4	
Extreme Muscle Rigidity	5		Facial muscles contorted and grimacing	5	

B) **Five Point Activity Scale For Sedation** per *ROSEN et al.*[44]

Score	Level of Activity
1	Comatose; does not respond to stimulation
2	Asleep; awakens with stimulation
3	Calm; catheters not at risk
4	Fussy; catheters at risk
5	Wild; no control, thrashing

C) **Modified Glasgow Coma Scale** per *JENNETTE et al.*[45]

EYE OPENING	**Score**
Spontaneous	4
To speech	3
To pain	2
None	1
BEST VERBAL RESPONSE	
Oriented	5
Confused	4
Inappropriate words	3
Incomprehensible sounds	2
None	1
BEST MOTOR RESPONSE	
Obeys	6
Localizes	5
Withdraws	4
Abnormal flexion	3
Extensor response	2
None	1

D) **SIMPSON AND REILLY SCALE**[46]

EYE OPENING	**SCORE**
Spontaneous	4
To speech	3
To pain	2
None	1
BEST VERBAL RESPONSE	
Oriented	5
Words	4
Vocal Sounds	3
Cries	2
None	1
BEST MOTOR RESPONSE	
Obeys commands	5
Localize pain	4
Flexion to pain	3
Extension to pain	2
None	1
NORMAL AGGREGATE SCORE	
Birth to 6 months	9
>6-12 months	11
>1-2 years	12
>2-5 years	13
>5 years	14

E) **The Children's Coma Score** per *RAIMONDI et al.*[47]

OCULAR RESPONSE	**SCORE**
Pursuit	4
Extraocular movements intact, pupils reactive	3
Extraocular movements impaired or pupils fixed	2
Extraocular movements paralyzed and pupils fixed	1
VERBAL RESPONSE	
Cries	3
Spontaneous respirations	2
Apneic	1
MOTOR RESPONSE	
Flexes and extends	4
Withdraws from painful stimuli	3
Hypertonic	2
Flaccid	1

F) **The O to IV Scale** per *SESHIA et al.*[48]

Arouses spontaneously and to stimuli	.O
Stuporous;spontaneous arousal rare;roused readily but briefly by stimuli; cough/gag present	I
Spontaneous arousal absent;semipurposive/avoidance motor response to stimuli; cough/gag depressed	II
Arousal in form of motor response only to intense,sustained, painful stimuli;cough/gag absent	III
Not aroused by even intense, sustained, painful stimuli; cough /gag absent	IV

G) **The Children's Orthopedic Hospital and Medical Center Scale** per *MORRAY et al.*[49]

CORTICAL FUNCTION	**SCORE**
Purposeful, spontaneous movements	6
Purposeful voice-evoked movements	5
Painful stimuli localized	4
Nonpurposeful movements/global withdrawal	3
Decorticate posturing	2
Decerebrate posturing	1
Flaccidity	0
BRAIN-STEM FUNCTION	
Intact	3
Depressed	2
Absent but breathing spontaneously	1
Absent and apneic	0

H) **Jacobi Scale** per *GORDON et al.*[50]

BEST VERBAL RESPONSE	**SCORE**
Fixes, follows, recognizes, laughs	5
Fixes and follows inconstantly, recognition uncertain	4
Arouses at times	3
Motor restlessness, unarousable	2
Complete unresponsiveness	1
BEST MOTOR RESPONSE	
Obeys	5
Localizes	4
Flexion	3
Extension	2
None	1
OCULOVESTIBULAR RESPONSE	
Normal	5
Tonic-conjugate	4
Minimal dysconjugate	3
No eye movements	2
Nonreactive pupil	1

I) **The Vancouver Sedative Recovery Scale** per *MACNAB et al.*[51]

RESPONSE		**SCORING OPTIONS**	***MOVEMENT***		**SCORING OPTIONS**
A.	i) Awake/alert	4	**I.**	i) Spontaneous and varied central activity	4
	ii) Awake/drowsy	3		ii) Spontaneous and varied peripheral activity	3
	iii) Asleep/easily aroused	2		iii) Central activity in response to stimuli	2
	iv) Asleep/difficult to arouse	1		iv) Peripheral activity in response to stimuli	1
	v) Asleep/unable to rouse	0		v) No movement	0
NOTE: IF CHILD SCORES (0) ON ABOVE , DO NOT PROCEED					
B.	i) Responds fully to stimuli in age-appropriate manner	2	**J.**	i) Absence of tremor or ataxia	1
	ii) Delayed response to stimuli	1		ii) Ataxia or tremor on being moved	0
	iii) Absent response to stimuli	0	**K**	i) Coordinated spontaneous movement	2
C.	i) "Alert" facial expression	1		ii) Weak/coarse spontaneous movement	1
	ii) "Flat" facial expression	0		iii) No purposeful spontaneous movement	0

I) **The Vancouver Sedative Recovery Scale** per *MACNAB et al.* continued[51]

	SCORING OPTIONS	*MOVEMENT*	SCORING OPTIONS
EYES		L. i)Shows age-appropriate manual dexterity	2
D. i) Bright eyes	1	ii)Awkward or clumsy hand movement	1
ii) Dull eyes; glazed	0	iii)No fine hand movement	0
E. i) Looks "at you"	1		
ii) Looks "through you"	0		
F. i) Accommodates	2		
ii) No attempt to accomodate	1		
iii) Unable to accomodate	0		
G. i) Recognition of stimulus	1		
ii) Limited or no recognition of stimulus	0		
H. i) Purposeful and spontaneous eye movement	1		
ii) Little or no spontaneous or purposeful eye movement	0		

TABLE 4

Goals of Sedation			
Prevent patient from pulling out IV lines/tubes			
Prevent self extubation by patient			
Promote patient safety through reduction of destructive movement by patient			
Promote Health Care Worker safety through reduction of destructive movement by patient			
Facilitate the execution of procedures(i.e. turning patient, bronchoscopy, line insertion)			
Prevent delirium tremors for patients undergoing alcohol withdrawal			
Reduce patient coughing episodes			
Enhance cooperation of patient to facilitate nursing care.			
Reduce nursing care demands during patient care			
Reduce intracranial pressure			
Optimize ventilation in patients who are asynchronous with the ventilator			
Reduce myocardial oxygen demand			
Optimize patients vital signs such as: heart rate, blood pressure, respiratory rate			
Reduce caloric expenditure			
Optimize pulmonary hemodynamics			
Optimize patients hemodynamic profile by reducing cardiac workload			

TABLE 4 CONTINUED

Goals of Sedation			
Induce short term amnesia			
Comply with patients request for sedation			
Reduce stress of patient of arousal			
Decrease level of consciousness in a paralyzed patient			
Reduce Aggression			
Reduce Anxiety			
Reduce Hallucinations			
Reduce Psychosis			
Reduce the amount of other medications (i.e. Antihypertensives) needed to optimize patients vital signs & decrease potential for their adverse effects			

TABLE 5

Please circle: Profession: M.D., Nurse, Pharmacist, Respiratory Therapist

Please assess and place each of the following Goals of Sedation in the appropriate classification with regards to the following statement:
** Clinically relevant and absolutely mandatory to ensure safe patient care**

Goals of Sedation	High Priority	Moderate Priority	Low Priority
Prevent patient from pulling out IV lines/tubes			
Prevent self extubation by patient			
Promote patient safety through reduction of destructive movement by patient			
Promote Health Care Worker safety through reduction of destructive movement by patient			
Facilitate the execution of procedures(i.e.turning patient,bronchoscopy,line insertion)			
Prevent delirium tremors for patients undergoing alcohol withdrawal			
Reduce patient coughing episodes			
Enhance cooperation of patient to facilitate nursing care.			
Reduce nursing care demands during patient care			
Reduce intracranial pressure			
Optimize ventilation in patients who are asynchronous with the ventilator			
Reduce myocardial oxygen demand			
Optimize patients vital signs such as: heart rate, blood pressure, respiratory rate			
Reduce caloric expenditure			
Optimize pulmonary hemodynamics			
Optimize patients hemodynamic profile by reducing cardiac workload			

TABLE 5 CONTINUED

Goals of Sedation	High Priority	Moderate Priority	Low Priority
Promote patient comfort by reducing pain			
Promote comfort to patient during procedures			
Promote regular sleep/wake cycle for patient			
Reduce family concerns by promoting patient comfort			
Induce short term amnesia			
Comply with patients request for sedation			
Reduce stress of patient on arousal			
Decrease level of consciousness in a paralyzed patient			
Reduce Aggression			
Reduce Anxiety			
Reduce Hallucinations			
Reduce Psychosis			
Reduce the amount of other medications (i.e. Antihypertensives) needed to optimize patients vital sign & decrease potential for adverse effects			

TABLE 6

Rank	Goal #	Goal
1	#24	Decreased level of consciousness in a paralyzed patient
2	#17	Promote patient comfort by reducing pain
3	#10	Reduce intracranial pressure
4	#11 & #2	Prevent self extubation by patient. Optimize ventilation in patients who are asynchronous with the ventilator
5	#18	Promote comfort to patient during procedures
6	#5	Facilitate the execution of procedures
7	#12	Reduce myocardial oxygen demand
8	#6	Prevent delirium tremors for patients undergoing alcohol withdrawal
9	#1 & #3	Prevent patient from pulling out IV lines/tubes. Promote patient safety through reduction of destructive movement by patient
10	#16	Optimize patients hemodynamic profile by reducing cardiac workload
11	#28	Reduce psychosis
12	#4 & #15	Promote health care worker safety through reduction of destructive movement by patient. Optimize pulmonary hemodynamics
13	#13	Optimize patients vital signs such as heart rate etc.
14	#25	Reduce aggression

The above goals are ranked in decreasing order of importance, based on the percentage of votes for high and medium priority. The goals ranked 1 -10 were considered high priority for greater than 50% of the votes.

TABLE 7

BEHAVIOR

Goal #1 - Prevent patient from pulling out IV lines/tubes.

Goal #3 - Promote patient safety through reduction of destructive movement by patient.

Goal #25 - Reduce aggression.

Goal #4 - Promote Health Care Worker safety through reduction of destructive movement by patient.

Goal #2 - Prevent self extubation by patient.

COMFORT

Goal #18 - Promote comfort to patient during procedures.

Goal #5 - Facilitate the execution of procedures(i.e. turning patient,bronchoscopy,line insertion).

Goal #17 - Promote patient comfort by reducing pain(i.e. sedation as adjuvant therapy to analgesics during burn dressing changes).

Goal #22 - Comply with patients request for sedation.

Goal #26 - Reduce anxiety.

Goal #23 - Reduce stress of patient on arousal.

Goal #7 - Reduce patient coughing episodes.

Goal #19 - Promote regular sleep/ wake cycle for patient.

Goal #8 - Enhance cooperation of patient to facilitate nursing care.

Goal #21 - Induce short term amnesia.

TABLE 7 CONTINUED

THERAPEUTIC

Goal #24 - Decrease level of consciousness in a paralyzed patient.

Goal #10 - Reduce intracranial pressure.

Goal #12 - Reduce myocardial oxygen demand.

Goal #11 - Optimize ventilation in patients who are asynchronous with the ventilator.

Goal #16 - Optimize patients hemodynamic profile by reducing cardiac workload.

Goal #13 - Optimize patients vital signs such as: heart rate,blood pressure,respiratory rate.

Goal #15 - Optimize pulmonary hemodynamics.

Goal #14 - Reduce caloric expenditure.

Goal #6 - Prevent delirium tremors for patients undergoing alcohol withdrawal.

Goal #29 - Reduce the amount of other medications(i.e.antihypertensives)needed to optimize patients vital signs & decrease potential for their adverse effects.

Goal #27 - Reduce hallucinations.

Goal #28 - Reduce psychosis.

TABLE 8

Goal #11 Optimize ventilation in patients who are asynchronous with the ventilator.

Oxygen Saturation

1) ≤ 85%

2) 86 - 90%

3) 91 - 95%

4) > 95%

Degree of Asynchrony	Patients Color
1) ≥75% asynchronous breaths	1) Waxy purple\blue face,lips and ears.
2) 26 - 74% asynchronous breaths	2) Red Blushing as during exercise
3) 1 - 25% asynchronous breaths	3) Pink or patients normal
4) 0% asynchronous breaths	

Paradoxical Breathing

1) Paradoxical breathing and accessory muscle use for a period of greater than 5 minutes.

2) Paradoxical breathing and accessory muscle use for a period lasting 3 to 5 minutes.

3) Paradoxical breathing and accessory muscle use for a period lasting >0 to 3 minutes.

4) No paradoxical breathing or excessory muscle use.

**ASYNCHRONY .....State where the patient and ventilator breathing cycles are mismatched.

**PARADOXICAL BREATHING....State where abdomen and chest are moving in opposite directions with accessory muscle (sternocleidomastoideus) use.

TABLE 9

LEVEL OF CONSCIOUSNESS

#1 Comatose, no response to painful stimuli or loud auditory stimulus, unable to rouse, patient does not obey commands and no eye opening present, no movement.

#2 Painful stimuli and/or loud auditory stimulus may rouse patient with difficulty or illicit a reflex response only, eyes may open, patient does not obey commands and has no purposeful movement.

#3 Painful stimuli and/or loud auditory stimulus cause patient to rouse and open eyes, patient has purposeful movement and may obey commands.

#4 Patient opens eyes spontaneously or via light touch or soft auditory stimulus, patient has purposeful movement and may obey commands.

*Note: Bedside nurses were trained to perform a standardized assessment of whether a patient can "Obey Commands". Patients were asked by bedside nurse to "open your mouth" or "stick out your tongue".

*Note: Bedside nurses were trained to perform standardized methods for patient rousal. The sequence in order was as follows: patient rouses spontaneously, if this is not the case nurse will try to rouse patient via soft voice and light touch; if this is still unsuccessful, nurse will try to rouse patient via a loud voice and light shaking of arm; if still unsuccessful, bedside nurse uses a loud voice with painful stimuli.

*Note: **Painful Stimuli....**will be firstly defined as nail bed pressure. If this painful stimuli can not be obtained in the patient, then sternal rub will be the second alternative.

TABLE 10

DESTRUCTIVE MOVEMENT

Goal #1 --Prevent patient from pulling out IV lines\tubes.

Goal #2---Prevent self extubation by patient.

Goal #3---Promote patient safety through reduction of destructive movement by patient.

#1 Continuous imminent risk of inducing harm to themselves or others or has produced harm to themselves or others and continues to be at risk of inducing harm(i.e. pulling out lines, movement out of bed etc.), fails to obey commands, 2 or more episodes within 1 hour where each episode takes greater than 1 minute for the nurse to calm patient.

#2 Intermittent episodes where patient is at imminent risk of harm to themselves or others.Patient disobeys commands due to forgetfulness of nursing commands, 1 episode every hour where each episode takes 1 minute or greater for nurse to calm patient.

#3 Periodic restlessness. Patient at minimal risk to harm themselves only. Patient obeys commands but forgets these nursing commands with time resulting in 2 to 3 episodes per day where each episode takes less than 1 minute for nurse to calm patient.

#4 Patient obeys commands and is tranquil. No risk of harm to themselves or others.

TABLE 11 **COMFORT**

Goal #18 - Promote comfort to patient during procedures

Goal #5 - Facilitate the execution of procedures (i.e. turning patient, bronchoscopy, line insertion).

Goal #17 - Promote patient comfort by reducing pain (i.e. sedation as adjuvant therapy to analgesics during burn dressing changes).

Goal #26 - Reduce anxiety

Degree of facial grimacing

#1 Facial grimacing constant with or without the patient undergoing a procedure.

#2 Facial grimacing occurs only during movement of patient during a procedure but does not occur when patient is at rest.

#3 No facial grimacing present during movement of patient during a procedure or at rest.

Voice and Movement Gestures

1) Ask the patient (whether intubated or not) if they are in any discomfort. **Patient must be awake and oriented.**

2) Patient will nod yes or no if in any discomfort or will try to "mouth" a response. If response is "yes" do a visual analogue scale to assess degree of discomfort. If no stop here.

"YES" 0----------10

**NOTE: The current standard visual analogue scale for pain in Adult ICU will be utilized for this assessment.

Patient can try to "mouth" response or will hold up fingers to indicate degree of discomfort on the scale of 1 to 10. If patient is a burn patient for example or has no finger dexterity then nurses can hold up their fingers and patient will nod to the number that corresponds to the degree of discomfort.

TABLE 12b **DISCUSSION AREA**

OXYGEN SATURATION

DEGREE OF ASYNCHRONY

PARADOXICAL BREATHING

LEVEL OF CONSCIOUSNESS

DESTRUCTIVE MOVEMENT

DEGREE OF FACIAL GRIMACING

VOICE AND MOVEMENT GESTURES

TABLE#12a	1 LEVEL	>1 LEVEL*	NO LEVELS*	DISCUSS*
Oxygen Saturation #1 ≤ 85% #2 86 - 90% #3 91 - 95% #4 > 95%				
Degree of Asynchrony #1 ≥75% asynchronous breaths #2 26 - 74% asynchronous breaths #3 1 - 25% asynchronous breaths #4 0% asynchronous breaths				
Patients Color #1 Waxy purple\blue face, lips and ears #2 Red Blushing as during exercise #3 Pink or patients normal				
Paradoxical Breathing #1 Paradoxical breathing and accessory muscle use for a period of greater than 5 minutes. #2 Paradoxical breathing and accessory muscle use for a period lasting 3 to 5 minutes. #3 Paradoxical breathing and accessory muscle use for a period lasting >0 to 3 minutes. #4 No paradoxical breathing or accessory muscle use.				
Level of Consciousness #1 Comatose, no response to painful stimuli or loud auditory stimulus, unable to rouse, patient does not obey commands and no eye opening present, no movement. #2 Painful stimuli and/or loud auditory stimulus may rouse patient with difficulty or illicit a reflex response only, eyes may open, patient does not obey commands and has no purposeful movement. #3 Painful stimuli and/or loud auditory stimulus cause patient to rouse and open eyes, patient has purposeful movement and may obey commands. #4 Patient opens eyes spontaneously or via light touch or soft auditory stimulus, patient has purposeful movement and may obey commands.				
Destructive Movement #1 Continuous imminent risk of inducing harm to themselves or others or has produced harm to themselves or others and continues to be at risk of inducing harm (eg. pulling out lines, movement out of bed etc.), fails to obey commands, 2 or more episodes within 1 hour where each episode takes greater than 1 minute for the nurse to calm patient. #2 Intermittent episodes where patient is at imminent risk of harm to themselves or others. Patient disobeys commands due to forgetfullness of nursing commands, 1 episode every hour where each episode takes 1 minute or greater for nurse to calm patient. #3 Periodic restlessness. Patient at minimal risk to harm themselves only. Patient obeys commands but forgets these nursing commands with time resulting in 2 to 3 episodes per day where each episode takes less than 1 minute for nurse to calm patient. #4 Patient obeys commands and is tranquil. No risk of harm to themselves or others.				
Degree of facial grimacing #1 Facial grimacing constant with or without the patient undergoing a procedure. #2 Facial grimacing occurs only during movement of patient during a procedure but does not occur when patient is at rest. #3 No facial grimacing present during movement of patient during a procedure or at rest.				
Voice and Movement Gestures 1) Ask the patient (whether intubated or not) if they are in any discomfort. Patient must be awake and oriented. 2) Patient will nod yes or no if in any discomfort or will try to "mouth" a response. If response is "yes" do a visual analogue scale to assess degree of discomfort. If no stop here. "YES" 0_____10 Patient can try to "mouth" response or will hold up fingers to indicate degree of discomfort on the scale of 1 to 10.				

TABLE 13a

Date:

Patient #:

Observer:

Oxygen Saturation

#1 ≤ 85%

#2 86 - 90%

#3 91 - 95%

#4 > 95%

Degree of Asynchrony (over a 5 minute period)

#1 ≥75% asynchronous breaths

#2 26 - 74% asynchronous breaths

#3 1 - 25% asynchronous breaths

#4 0% asynchronous breaths

Patients Color

#1 Waxy purple\blue face, lips and ears

#2 Red Blushing as during exercise

#3 Pink or patients normals

Paradoxical Breathing (over a 5 minute period).

#1 Paradoxical breathing for a period of greater than 5 minutes.

#2 Paradoxical breathing for a period lasting 3 to 5 minutes.

#3 Paradoxical breathing for a period lasting >0 to< 3 minutes.

#4 No paradoxical breathing

Accessory Muscle Use (over a 5 minute period).

#1 Accessory muscle use for a period of greater than 5 minutes.

#2 Accessory muscle use for a period lasting 3 to 5 minutes.

#3 Accessory muscle use for a period lasting >0 to <3 minutes.

#4 No accessory muscle use.

Level of Consciousness

#1 Comatose, no response to painful stimuli or loud auditory stimulus, unable to rouse, patient does not obey commands and no eye opening present, no movement.

#2 Painful stimuli and/or loud auditory stimulus may rouse patient with difficulty or illicit a reflex response only, eyes may open, patient does not obey commands and has no purposeful movement.

#3 Painful stimuli and/or loud auditory stimulus cause patient to rouse and open eyes, patient has purposeful movement and may obey commands.

#4 Patient opens eyes spontaneously or via light touch or soft auditory stimulus, patient has purposeful movement and may obey commands.

Degree of facial grimacing

#1 Facial grimacing at rest.

#2 Facial grimacing occurs only during manipulation but does not occur when patient is at rest.

#3 No facial grimacing present(i.e. at rest or manipulation).

Voice and Movement Gestures

1) Ask the patient (whether intubated or not) if they are in any discomfort. **Patient must be awake, oriented.**

2) Patient will nod yes or no if in any discomfort or will try to "mouth" a response. If response is "yes" do a visual analogue scale to assess degree of discomfort. If no stop here. "YES" 0_____10

Patient can try to "mouth" response or will hold up fingers to indicate degree of discomfort on the scale of 1 to 10.

DEFINITIONS:

Episode: is defined as the state where a patient is at risk of inducing harm to themselves or others or has produced harm to themselves or others and continues to be at risk of inducing harm.

Oriented: is defined as the state where the patient can follow simple commands and able to comprehend the language spoken.

Asynchrony: is defined as the state where the patient and ventilator breathing cycles are mismatched.

Paradoxical Breathing: is defined as the state where abdomen and chest are moving in opposite directions with accessory muscle(sternocleidomastoideus) use.

Painful Stimuli: will be administered by bedside nurse and will consist of nail bed pressure first followed by sternal rub as a secondary alternative, if applicable.

TABLE 13b

DATE:
PATIENT#:

Primary Caregiver:

Destructive Movement

#1 - 2 or more episodes within 1 hour where each episode takes greater than 1 minute for the nurse to calm patient.

#2 - 1 episode every hour where each episode takes 1 minutes or greater for nurse to calm patient.

#3 - 6 or less episodes per 12 hour shift, where at least one episode takes greater than 1 minute for nurse to calm patient.

#4 - 6 or less episodes per 12 hour shift, where each episode takes less than 1 minute for nurse to calm patient.

#5 - Patient is tranquil. No risk of harm to themselves or others.

Definitions

Episode: is defined as the state where a patient is at risk on inducing harm to themselves or other or has produced harm to themselves and continues to be at risk of inducing harm.

TABLE 13c

DATE:
PATIENT#:

Secondary Caregiver:

Destructive Movement

#1 - 2 or more episodes within 1 hour where each episode takes greater than 1 minute for the nurse to calm patient.

#2 - 1 episode every hour where each episode takes 1 minutes or greater for nurse to calm patient.

#3 - 6 or less episodes per 12 hour shift, where at least one episode takes greater than 1 minute for nurse to calm patient.

#4 - 6 or less episodes per 12 hour shift, where each episode takes less than 1 minute for nurse to calm patient.

#5 - Patient is tranquil. No risk of harm to themselves or others.

Definitions:

Episode: is defined as the state where a patient is at risk of inducing harm to themselves or others or has produced harm to themselves and continues to be at risk of inducing harm.

TABLE 14 Weighted k (kappa) Summary

k (kappa) Value	Interobserver Agreement
Less than 0	Poor
0 to 0.20	Slight
0.21 to 0.40	Fair
0.41 to 0.60	Moderate
0.61 to 0.80	Substantial
0.81 to 1.00	Almost Perfect

TABLE 15

				Distribution of Data Among Levels (%)				
Monitoring Parameter	% Agreement	Disagreement Rate	Weighted Kappa	Level #1	Level #2	Level #3	Level #4	Level #5
Destructive Movement	68.72	0.3128	0.3889	10.66	6.87	10.90	21.09	50.47
Oxygen Saturation	99.12	0.0088	0.9553	0	2.08	24.40	73.51	
Degree of Asynchrony	97.05	0.0295	0.3062	0.61	0	3.87	95.36	
Color	97.93	0.0207	0.2015	0.89	1.19	97.91		
Paradoxical Breathing	93.21	0.0679	0.1857	4.46	1.19	2.08	92.26	
Accessory Muscle Use	87.59	0.1241	0.2388	6.25	6.55	2.98	84.23	
Level of Consciousness	90.54	0.0946	0.6719	7.44	13.99	7.74	70.83	
Degree of Facial Grimacing	92.92	0.0708	0.6858	5.06	22.92	72.02		
Voice and Movement Gestures	99.71	0.0029	0.9873	37.80	62.20			

Overall % Agreement 91.87%
Overall Disagreement Rate 0.0813
Overall Kappa 0.5137

Bibliography

1) Crippen DW. The Role of Sedation in the ICU Patient with Pain and Agitation. CRITICAL CARE CLINICS 1990;6(2):369-392.

2) Blount RL, Corbin SM, Sturges JW, Wolfe VV, Prater JM, James LD. The Relationship Between Adults' Behavior and Child Coping and Distress During BMA/LP Procedures: A Sequential Analysis. BEHAVIOR THERAPY 1989;20:585-601.

3) Katz ER, Kellerman J, Siegel SE. Behavioral Distress in Children Undergoing Medical Procedures: Developmental Consideration. JOURNAL OF CONSULTING AND CLINICAL PSYCHOLOGY 1980;48:356-365.

4) Marx CM, Rosenberg DI, Ambuel B, Hamlett KW, Blumer JL. Survey of Sedative Utilization in Pediatric Critical Care. 3RD PEDIATRIC CRITICAL CARE COLLOQUIUM 1989;October: Santa Monica CA.

5) Aitkenhead AR. Analgesia and Sedation in Intensive Care. BR. J. ANAESTH 1989;63:196-206.

6) Xiaolu S, Quinn T, Weissman C. Patterns of Sedation and Analgesia in the Postoperative ICU Patient*. CHEST 1992;101(6):1625-1632.

7) Adam K, Oswald I. Sleep Helps Healing. BRITISH MEDICAL JOURNAL 1984;28(Nov):1400-1401.

8) Merriman HM. The Techniques Used to Sedate Ventilated Patients. INTENSIVE CARE MEDICINE 1981;7:217-224.

9) Halmagy DFJ, Kinney JM. Metabolic Rate in Acute Respiratory Failure Complicating Surgery. SURGERY 1975;77:492-497.

10) Ritz R. Benzodiazepine Sedation in Adult ICU Patients. INTENSIVE CARE MEDICINE 1991;17:S11-S14.

11) Rapold HJ, Follath F, Scollo-Lavizzari G, Ritz R. Diazepam: Hazardous Sedation of Ventilated Patients. INTENSIVE CARE MEDICINE 1983;9:218.

12) Burns AM, Shelly MP, Park GR. The Use of Sedative Agents in Critically Ill Patients. DRUGS1992;43(4):507-515.

13) Bion JF, Ledingham I McA. Sedation in the Intensive Care - A Postal Survey. INTENSIVE CARE MEDICINE 1987;13:215-216.

14) Marx CM, Smith PG, Lowrie LH, et al. Optimal Sedation of Mechanically Ventilated Pediatric Critical Care Patients. CRITICAL CARE MEDICINE 1994;22(1):163-170.

15) Kong KL, Willatts SM, Prys-Roberts C. Isoflurane Compared with Midazolam for Sedation in the Intensive Care Unit. BRITISH MEDICAL JOURNAL 1989;298(May):1277-1280.

16) Bion JF, Logan BK, Newman PM, Brodie MJ, Oliver, JS, Aitchison TC, Ledingham I McA. Sedation in Intensive Care: Morphine and Renal Function. INTENSIVE CARE MEDICINE 1986;12:359-365.

17) The Shorter Oxford English Dictionary. OXFORD UNIVERSITY PRESS 1986;2:1927.

18) O'Sullivan GF, Park GR. The Assessment of Sedation in Critically Ill Patients. JOURNAL OF CLINICAL INTENSIVE CARE 1990;1:116-121.

19) Hansen-Flaschen J, Cowen J, Polomano RC. Beyond the Ramsay Scale: Need for a Validated Measure of Sedating Drug Efficacy in the Intensive Care Unit. CRITICAL CARE MEDICINE 1994;22(5):732-733.

20) Ramsay M, Savage T, Simpson B, et al. Controlled Sedation with Alphaxolone-Alphadolone. BR MED J 1974;2:656-659.

21) Cook S, Palma O. Propofol As A Sole Agent For Prolonged Infusion in Intensive Care. JOURNAL OF DRUG DEVELOPMENT 1989;2(Suppl. 2):65-67.

22) Hopkinson RB, Freeman JW. Therapeutic Progress - Review XXX: Therapeutic Progress in Intensive Care Sedation and Analgesia: Part I - Principles. JOURNAL OF CLINICAL PHARMACY AND THERAPEUTICS 1988;13:33-40.

23) Grounds RM, Lalor JM, Lumley J, Royston D, Morgan M. Propofol Infusion for Sedation in the Intensive Care Unit: Preliminary Report. BRITISH MEDICAL JOURNAL 1987;294(Feb):397-400.

24) Taub HA, Eisenberg L. An Evaluation of Memory Under Regional Anesthesia with IV Lorazepam As A Premedicant. ANESTH ANALG 1976;55:368-73.

25) Samra SK, Bradshaw EG, Pandit SK, Papanicolaou AC, Moore BD. The Relation Between Lorazepam-Induced Auditory Amnesia and Auditory Evoked Potentials. ANESTH ANALG 1988;67:526-33.

26) Chernik DA, Gillings D, Laine H, Hendler J, Silver JM, Davidson AB, Schwam EM, Siegel JL. Validity and Reliability of the Observer's Assessment Of Alertness/Sedation Scale: Study With Intravenous Midazolam. JOURNAL OF CLINICAL PSYCHOPHARMACOLOGY 1990;10(No.4):244-251.

27) Teasdale G, Jennett B. Assessment of Coma and Impaired Consciousness. LANCET 1974;2:81-83.

28) Barvais L, Dejonckheere M, Dernovoi B. Continuous Infusion of Midazolam or Bolus of Diazepam for Postoperative Sedation of Cardiac Surgical Patients. ACTA ANAESTHIOL BELG 1988;39:239-245.

29) Carrasco G, Molina R, Costa J. Propofol vs. Midazolam in Short-, Medium-, and Long-Term Sedation of Critically Ill Patients: A Cost-Benefit Analysis. CHEST 1993;103:557-564.

30) Ellis A, Cavanagh SJ. Aspects of Neurosurgical Assessment Using the Glasgow Coma Scale. INTENSIVE CRIT. CARE NURS 1992;8(2):94-99.

31) Bhatty GB, Kapoor N. The Glasgow Coma Scale: A Mathematical Critique. ACTA NEUROCHIR WIEN 1993;120(3-4):132-135.

32) Rowley G, Fielding K. Reliability and Accuracy of the Glasgow Coma Scale with Experienced and Inexperienced Users. LANCET 1991;337(8740):535-538.

33) Stalhammar D, Starmark JE. Assessment of Responsiveness in Head Injured Patients. ACTA NEUROCHIR 1986;Suppl 36:91-94.

34) Teasdale G, Jennett B, Murray L, Murray G. Glasgow Coma Scale: To Sum or Not To Sum? LANCET 1983;2:678.

35) Menegazzi JJ, Davis EA, Sucov AN, Paris PM. Reliability of the Glasgow Coma Scale when used by Emergency Physicians and Paramedics. THE JOURNAL OF TRAUMA 1993;34(1):46-48.

36) Côté R, Battista R, Wolfson C, Boucher J, Adam J, Hachinski V. The Canadian Neurological Scale. NEUROLOGY 1989;39(Nov):1556-1558.

37) Giancino JT, Kezmarsky MA, DeLuca J, Cicerone KD. Monitoring Rate of Recovery to Predict Outcome in Minimally Responsive Patients. ARCH PHYS MED REHABIL 1991;72(Oct):897-901.

38) Coren S, Mah KB. Prediction of Physiological Arousability: A Validation of the Arousal Predisposition Scale. BEHAV RES. THER. 1993;31(2):215-219.

39) McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J, Chapman J. CHEOPS: A Behavioral Scale for Rating Postoperative Pain in Children. ADVANCES IN PAIN RESEARCH AND THERAPY 1985;9:395-402.

40) Whipple JK, Quebbeman EJ, Lewis KS, Gottlieb MS, Ausman RK. Difficulties in Diagnosing Narcotic Overdoses in Hospitalized Patients. THE ANNALS OF PHARMACOTHERAPY 1994;28(April):446-450.

41) Cataldo MF, Bessman CA, Parker LH, Reid Pearson JE, Rogers MC. Behavioral Assessment for Pediatric Intensive Care Units. JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1979;12(1):83-97.

42) Gall J, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, Mercier P, Thomas R, Villers D. A Simplified Acute Physiology Score for ICU Patients. CRITICAL CARE MEDICINE 1984;12(11):975-977.

43) Ambuel B, Hamlett KW, Marx CM, Blumer JL. Assessing Distress in Pediatric Intensive Care Environments: The Comfort Scale. JOURNAL OF PEDIATRIC PSYCHOLOGY 1992;17(1):95-109.

44) Rosen DA, Rosen KR. Midazolam for Sedation in the Pediatric Intensive Care Unit. INTENSIVE CARE MEDICINE 1991;17:S15-S19.

45) Jennett B, Teasdale G, Galbraith S, et al. Severe Head Injuries in Three Countries. J NEUROL NEUROSURG PSYCHIATRY 1977;40:291-298.

46) Simpson D, Reilly P. Paediatric Coma Scale. LANCET 1982;2:450.

47) Raimondi AJ, Hirshcaver J. Head Injury in the Infant and Toddler: Coma Scoring and Outcome Scale. CHILDS BRAIN 1984;11:12-35.

48) Seshia SS, Seshia MMK, Sachdeva RK. Coma in Childhood. DEV. MED CHILD NEUROL 1977;19:614-628.

49) Morray JP, Tyler DC, Jones TK, Stuntz JT, Lemire RJ. Coma Scale for use in Brain-Injured Children. CRITICAL CARE MEDICINE 1984;12:1018-1020.

50) Gordon NS, Fois A, Jacobi G, Minns RA, Seshia SS. Consensus Statement: The Management of the Comatose Child. NEUROPEDIATRICS 1983;14:3-5.

51) Macnab AJ, Levine M, Glick N, Susak L, Baker-Brown G. JOURNAL OF PEDIATRIC SURGERY 1991;26(11):1263-1267.

52) Teasdale G, Knill-Jones R, Van der Sande J. Observer Variability in Assessing Impaired Consciousness and Coma. J NEUROL NEUROSURG PSYCHIATRY 1978;41:603-610.

53) Feinstein AR. Clinimetrics. NEW HAVEN, CONN: YALE UNIVERSITY PRESS;1987.

54) Fleiss JL. STATISTICAL METHODS FOR RATES AND PROPORTIONS. New York, NY: John Wiley And Sons Inc;1981:212-236.

55) Cohen J. A coefficient of agreement for nominal scales. EDUC PSYCHOL MEAS 1960;20:37-46.

56) Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. BIOMETRICS 1977;33:159-174.

57) Rosner B. FUNDAMENTALS OF BIOSTATISTICS. Boston, Prindle, Weber, and Schmidt;1986:424-429.

58) Tyler DC, Douthit ATJ, Chapman CD. Toward Validation of Pain Measurement Tools for Children: A Pilot Study. PAIN 1993;52:301-309.

59) Gaddis ML, Gaddis GM. Introduction to Biostatistics: Part 6, Correlation and Regression. ANN-EMERG-MED 1990;19(12):1462-1468.

60) Hassard TH. UNDERSTANDING BIOSTATISTICS. St. Louis: Mosby-Year Book Inc., 1991:127-134 and 273-275.

61) Schefler, William C. (1984) Statistics for Health Professionals. Addison-Wesley, Reading, Massachusetts. pp. 257-260.

62) Shitzer RL, Cohen J, Fleiss JL, Endicott J. (1967) Quartification of Agreement in Psychiatric Diagnosis. ARCH GEN PSYCHIATRY 17:83-87.