

CHAPTER III

DISCUSSION AND SUMMARY

3.1 Discussion:

There is a variety of scales to assess the consciousness of the patients. However the simplicity and validity of the GCS has gained it widespread acceptance for evaluating neurologic function and been familiar to physicians, nurses and paramedics of all specialties, including neurosurgery(8). The GCS has been used not only to grade individual levels of consciousness, indicate prognosis, compare effectiveness of treatment but also improve clinical communication between medical personnel(26). Nevertheless, the limitation of GCS has been mentioned emphatically on the uncertainty in functional brainstem assessment and the scoring doubt in some circumstances; it has led to proposals for alternate ways of neurologic evaluation.

Debate about the inter-rater reliability and vague terms used to describe the level of consciousness of the GCS has been proposed. Teasdale et al had revealed disagreement rates as high as 0.191 among observers approaching a whole scale point(27). Subsequent researches have found the contradictions including our finding(1, 14, 28, 29). From the current observation, we found excellent reliability of the FOUR score and the GCS within group identifications no matter either limited training and experience or well-trained practitioners. Every classifications of evaluators; expert clinicians, novice clinician, experienced registrar nurse and apprenticed registrar nurse have tremendously high level of reliability in the GCS assessment likewise the FOUR score. Wijdicks et al had reported good to excellent inter-rater reliability in FOUR score assessment among neuroscience nurses, neurointensivists, and proficient neurology residents(7) as well as Wolf et al had reported in intensive care nurses similar to our finding(29). The result of extraordinary level of overall reliability is possibly caused by thoroughly training, providing standard definitions as guidance, and full demonstrations of the evaluation process in all examiners as mentioned in the methodology section.

The dwindling level of reliability in brainstem subscales for FOUR score and eye subscales for the GCS are noticeable. Ratings are influenced by several factors specifically the loudness of observer's voice, the intensity of pain stimulation, the estimation or even the measurement method for the pupil sizes, the time spent to obtain the response, or the fluctuating alertness in between rating. The Wilcoxon's signed rank test demonstrated the insignificant difference between the sorts of all examiners. Meticulously, the average score of both scales were substantially high. The consciousness stratification and practicability in extreme range of both dimensions may be exploited but the sensitivity in the intermediate range should be deliberated. Starmark et al found that the greatest difficulty of the users' assessment was with the middle of range scores which confirms the suggestion made by Rowley et al(30). They demonstrated that the inexperienced users, although maintaining an acceptable level of agreement, made substantial errors depending on the conditions of the patients observed, averaging up to one point on each subscale of the GCS(14).

Cronbach's alpha was calculated from expert clinicians' data to show the internal consistency of the FOUR score and the GCS. The value of 0.82 from the FOUR score and 0.85 from the GCS are considered good but not enough to be used as a research tool to compare groups. A value of 0.9 would be desirable(31). The discriminating ability of this consistency is not involved in the scoring essence because all components were maintained but the variability may be commenced. The definitions of impaired consciousness are often inconsistent and ambiguous. The consciousness categorization from the beginning was somewhat given erratic results by the gradient of the physician screening decision across time and situations. The intra-rater reliability should be augmented to minimize this effect.

An assortment of validity precisely clinical, construct and criterion concern was appraised. The FOUR score also has potential capability to classify state of consciousness and predict mortality risk among the neurosurgical patients similar to the GCS indicated by the preeminent value of Spearman's rank correlation coefficient. Several reports established the relationship of the consciousness deterioration or improvement scale to the prediction of the outcomes(9, 11, 22-24). The endeavor to designate mortality risk therefore the eventual outcome; in-hospital mortality rate, 3-

month and 6-month mortality rate should be certainly assessed. Nonetheless, the time period of data collection was inappropriate; the GCS total score was then substituted in the balance of evidence-based deliberated previously by Bastos et al(21). In general, the higher the FOUR score, the better the outcome. Thus, the total FOUR score of 0-7 has been indicated high mortality risk while score 8-14 and 15-16 determined intermediate and low mortality risk subsequently from this study.

Another endorsement of consciousness evaluation scales comparing with the GCS was made in 1981. Salzman et al asserted that The Maryland Coma Scale score of 35% or less was used as the cut off point for a definition of deep coma or severe head injury compared to the GCS score equal or less than 7(4). The prognostications of final outcomes, e.g. mortality rate, the percentage of recovery divided by the period of observation in days which represented "recovery rate", etc. based on the initial score can also be predicted from this statistical analysis.

Even the mortality rate between the groups of patients can be estimated from FOUR score, the decisiveness should always be used in association with information about the patient's primary diagnosis, operative condition, age and chronic health status, since survival may be affected by the differing pathology. It should be recognized that the FOUR score is able to differentiate patients likely to live or dead, its discriminating ability should be considered vigorously. The aggressive therapies and interventions can alter an acceptable mortality.

The practical feasibility among raters reached statistical difference between the FOUR score and the GCS, as denoted by the p value. This could be a consequence of teaching the GCS in medical related curriculum and its frequent use in intensive care units and other areas of the hospital(29). Although the lower practicable score of the FOUR score, it has met all the requirements of a practical, investigative, yet psychometrically sound, clinical instrument. Following a period of familiarization with this scale, the FOUR score can be administered in the appropriate amount of time required and interpreted indistinguishable with the GCS.

The apparent variation in the use of some medications emphasizes the substantial risk involved in inferring a neurologic assessment(21). In accordance with this statement, the patients who were receiving sedatives, hypnotics, neuromuscular

blocking agents, or recovering from general anesthesia were excluded. Unfortunately, most of the consciousness evaluating score is routinely used in these situations, but the regulation was prepared. It is assumed that controlling for these factors might improve the evaluation of unadulterated effect and result in outcome predictions from neurologic feature for the most part(8). Homer et al has reported the prognostic significance of the alertness scale is reduced if one of the components is disturbed by the surrounding milieu(11). The application of painful stimuli was another precaution. Any pressure on the supraorbital ridge, sternum or nail beds may induce a spinal reflex response(8). Only a pinch to the pectoralis major muscle, an acceptable type of central painful stimulus was allowed.

Comparisons between the reliability measurements obtained from different patients may not be valid if it is likely that their conditions differ substantially(14). In this current study, the tasks presented to the four groups of evaluators were equal difficulty, ensued the stratified random sampling of the patients in each consciousness category.

Following the research objectives, the author has proposed that the FOUR score is reliable among the various assessor classifications, valid enough for consciousness evaluation and powerful for the estimation of mortality risk in neurosurgical patients. Further study of consciousness assessment in emergency situation, traumatic brain injury and pediatric patients should be done to verify the foregone conclusion of the FOUR score as well as the investigation of mortality rate, in order to certify the cut points estimation of the assorted risk directly, not using the surrogate outcome as described limitation.

3.2 Conclusion:

Considerable experience with the GCS has validated its usefulness as a prognostic instrument and its importance in standardizing cross study reporting for the purpose of meaningful comparison. Nevertheless, in the evaluation of new scoring modalities, a research tool with a finer grain may well be required, especially when performing serial evaluations. The ideal conscious grading system should be based on objective criteria that clearly delineate each progression, reproducible, represent a

continuum of clinical conditions from the alertness to the comatose patient, compare effectiveness of treatment and empower to define and predict patient outcome.

The FOUR score was entirely assembled the criteria remarkably in neurosurgical patients. This scale remains testable in intubating patients, recognizes brainstem reflexes and provides information about the presence of a respiratory drive. It has provided a standardized approach that health care practitioners can utilize to monitor and trend neurologic assessment findings in the absence of expensive or invasive techniques, reliability and consistence when evaluating the clinical effects. The lower practicability score of the FOUR score compared with the GCS was found but the universal implementation may provide the opportunity for the better clinical practice and resulting high quality of patient care.

3.3 Limitation:

There were several limitations to this study. An observation could introduce a variety of biases including those related to ascertainment. Test-retest reliability had not been evaluated because it was impossible for the rater to score the same patient in such an abrupt time period without remembering the previous scores. Hopefully, the influence of these biases will be reduced by the focus on eligibility criteria and objective patient characteristics.

For the circumstance of outcome prediction, the in-hospital mortality should be followed and recorded. Because of the time limited to this study, the author depicted the outcome prediction derived from the GCS at 2 cut points as Bastos et al study instead of the actual mortality(21).

Additionally, the consciousness assessment with the FOUR score and the GCS in pediatric patient might be obviously inappropriate because of the deprivation of mental development. In practical situation, the application of these scales must be considered.

3.4 Ethical issues:

The consciousness evaluation is the routine practice among the neurosurgical patients but the caution must be used in applying painful stimulus. The author try to

diminish the possible harm from the assessment by designated only a pinch to the pectoralis major muscle not any pressure on the supraorbital ridge or the sternum while demonstrating painful stimuli, the usage of sterile cotton fiber or 2-3 sterile normal saline eye drops from the distance of 4-6 inches for corneal reflex evaluation and the allowance only tracheal suction to evaluate cough reflex. All of the procedures mentioned are clinically accepted(7, 8).

Informed consent must be signed in every single cases by the patients or the legally relatives. The ownership permission must be obtained whenever they can comprehend instantaneously. The written protocol to obtain patient data must be approved by the Ethical committee.

3.5 Economic consideration:

A model of cost-effectiveness or cost-minimization is concerned on a provider viewpoint but the decision making would depend on the outcomes. Nevertheless, both measurements are difficult to value the accurate cost especially on a practitioner time consumed to obtain the test. The author proposes that they assumed to be equal so that the economic matter would be omitted.