CHAPTER IV

The Seizure Severity Reliability Test in addition to Counting the Number of Seizure Attacks

Background and Rationale

In the current situation of general practice, all physicians assess the efficacy of treatment of epilepsy patients by using the magnitude of seizure frequency reduction or seizure free period. Some patients might have many types of seizure, which do not have the same features contributing to a difference in severity even though the total frequency of seizures is not reduced. Examples are the reduction in severity from secondary generalized tonic clonic convulsion (2 GTC) to complex partial seizure (CPS) or from complex partial seizure with altered consciousness to simple partial seizure (SPS) with consciousness maintained.

Regarding the features of seizure as well, the same type of seizure might have a difference in features that produce a different severity. For example, a

patient with generalized tonic clonic convulsion (GTC) has no clonic jerking of limbs after treatment or patient who has a shortening post ictal period resumes activity more rapidly. Therefore using the frequency of unspecified seizures measured as a unique variable to assess outcome could lead to vague and biased results.

As mentioned above, seizure frequency could be a useful index if the specificity of each seizure and the context of other seizure variables are taken into consideration. Hence, the types and features of seizure altogether will be essential to evaluate the efficacy of treatment.

Literature Review

Up to now, there have been four scales developed to quantify seizure severity from seizure feature. Before tools are borrowed from other fields, the following questions should be asked and answered:

- Is the measure applicable to answering questions with respect to epilepsy?
- Do we have standards against which the measures can be validated for people with epilepsy?
- Do we know the relationships and overlapping redundancies between the large variety of measurements used for outcome evaluation in epilepsy?
- Is their former validity useful?

- What is the relative value of the measures in particular as related to frequency assessment?

- Are they adapted for the particular cultural environment in which we intend to use them?

A review of reports on outcome measurement in epilepsy shows a wide variety of measures of seizure severity; on the one hand, using objective frequency events and on the other using subjective perceptions of seizure severity or a mixture of both. The severity of seizures and their clinical manifestations may have equal or greater impact on patient well being as the total number of seizures. The best scale should have high reliability, validity and sensitivity to change.

The first scale (Baker GA,... [et.al.]. (1998)) to measure the severity of seizure is Veterans Administration Seizure Frequency and Severity Rating Scale (VA). The VA scale was the first scale developed specifically for clinical trials to measure seizure severity. The rating quantifies the frequency and severity of seizures, based on scores assigned for GTC, CPS, and SPS. These scores are modified by a variety of factors frequently reported by patients as important in determining the severity of their seizures: sleep deprivation, warning/aura, nocturnal timing, fever or illness, and missed dose of antiepileptic drugs (AEDS). This scale was used in two VA multicenter studies but only a small study of interrater agreement is available. This scale is a major component of a global composite score when added to the systemic and neurotoxic effects of AEDS. The validity of this scale comes from patient assessment. The advantage of this scale is patients

report the severity of their seizures and utilization information comes from patients and witnesses. The disadvantage is having a small study of inter-rater agreement.

The second scale (Donoghue MF, Duncan JS and Sander JWAS. (1996)) is the Liverpool Seizure Severity Scale (LS). It has two main factors: one is patients' perception of control over their seizures called percept subscale, the other is severity of ictal and post ictal phenomena called ictal subscale. Patients rate major or minor seizures according to their feeling of those types of seizure occurrence. This scale is a self-administered questionnaire of 16 questions but the latest version has 20 questions consisting of 2 subscales, a 11-item of ictal scale and a 9-item of percept subscale. An expert panel chose those items (shown in Table VII). A 4-point Likert scale is applied to each item.

The reliability data have been presented in the form of a test-retest Pearson correlation coefficient of 0.8 for both sub-scales and of a Cronbach's alpha of 0.85 for the ictal subscale and 0.69 for the percept subscale which is below that which is acceptable. This percept sub-scale is also not sensitive to change.

The advantage of this scale is patients rate their major or minor seizures if they have more than one seizure type. As well, the reliability test and test of internal consistency of ictal subscale is sufficient. The disadvantage is the items were chosen by an expert panel. As well, test of internal consistency and sensitivity to change of percept subscale is unacceptable.

***Table VII:** Liverpool Seizure Severity Scale Items

Ictal items	Percept items
Falls	Presence of diurnal cycle
Urinary incontinence	Presence of aura
Tongue biting	Control over attacks
Other injuries	Clustering of seizures
Recovery time	Seizures during sleep
Postictal headache	Being able to tell when I will have attacks
Postictal sleepiness	Being able to fight off attacks
Postictal confusion	Interference with the things I want to do
Duration of confusion	Subjective severity of attacks
Loss of consciousness	
Automatisms	

* Source: Donoghue MF, Duncan JS and Sander JWAS. (1996). The national hospital seizure severity scale: A further development of the chalfont seizure severity scales. Epilepsia. (37) : 563-571.

The third scale (Duncan JS and Sander JWAS. (1991)) is Chalfont Seizure Severity Scale (CS). It consists of 11-item scale (shown in Table VIII) focusing solely on the objective clinical events of a seizure and is administered by interviewing a patient and a witness to the seizures. Like the LS, the scale was applied to different seizure types separately. The scale's content was derived from open interviews with people with epilepsy. Most items had a 5-point Likert scale-scoring system based on the frequency of occurrence of that item. The scoring of the scale was derived using a combination of patient and expert opinion to create an acceptable ranking of scores from different seizure types. Each item has a

difference in weighting score that came from patient's weighting and adjusted by repeated doing the scale. The reliability test of this scale has been tested by inter-rater and test-retest methods that have sufficient reliability. The face and content validity is present but the construct validity has not been reported. It may be responsive to change. The advantage of the scale is that it has sufficient levels of reliability and face and content validity and that the items came from patients and witnesses.

***Table VIII:** Chalfont Seizure Severity Scale Items

- Loss of consciousness
- Aura
- Dropping or spilling objects
- Falls
- Injuries
- Incontinence
- Automatisms
- Generalized convulsions
- Duration of seizure
- Duration of recovery time
- Seizures only in sleep

*Source: Donoghue MF, Duncan JS and Sander JWAS. (1996). The national hospital seizure severity scale: A further development of the chalfont seizure severity scales. Epilepsia. (37) : 563-571.

The last scale (Donoghue MF, Duncan JS and Sander JWAS. (1996)) is National Hospital Seizure Severity Scale (NH3). It was derived from CS by elimination of 4 redundant items, a change in content in one item and a simplification of the scoring system. The items eliminated are dropping objects because it paralleled the item on falls, seizures occurring only in sleep because change on this item was rarely observed, time to complete recovery was replaced be separating items on seizure duration and recovery phase, and loss of consciousness was incorporated into the question on aura. The item on injuries was changed from frequency to severity of injuries. The weighting scores from CS have been changed to simple score in NH3. The researcher did all of changed items (shown in Table IX). The construct validity was demonstrated by agreement of patients with the severity score of certain type of seizure created by researcher. Patients rank and rate the severity of the 5 created prototype seizures. The interobserver and test-retest reliability were done and reported with intraclass correlation coefficient (ICC). The ICC was 0.9. The advantage is score of items came from patients' rating and having sufficiency of reliability test. The disadvantage is researcher did all of changed items.

The CS was chosen because all items came from the patients' and close relatives' concern with acceptable reliability.

***Table IX:** The National Hospital Seizure Severity Scale (NH3)

1.Record the name of the seizure types that occur under headings "type 1, 2, 3..."

2. Does the patient have a generalized convulsion during this type of seizure?

Yes = score 4 No = score 0

3. How often has the patient fallen to the ground in this type of seizure?

Nearly always or always	=	score 4
Often	=	score 3
Occasionally	=	score 2
Never	=	score 0

4. Has this type of seizure caused any of the following? (score only the worst)

Burns, scalds, deep cuts, fractures	=	score 4
Bitten tongue or severe headache	=	score 3
Milder injuries or mild headache	=	score 2
No injuries	=	score 0

5. How often has the patient been incontinent of urine in this type of seizure?

Nearly always or always	=	score 4
Often	=	score 3
Occasionally	=	score 2
Never	=	score 0

6. If the seizure causes loss of consciousness, is there a warning long enough for the patient to protect him/herself? (no loss of consciousness or seizure only while asleep scores 0)

Never	=	score 2
Sometimes	=	score 1
Nearly always or always	=	score 0

7. How long is it until the patient is really back to normal after the seizure?

Less than 1 minute	=	score 0
Between 1 and 10 minutes	=	score 1
Between 10 minutes and 1 hour	=	score 2
Between 1 and 3 hours	=	score 3

More than 3 hours	=	score 4
8. Do the following events occur in this type of seizure?		
Seriously disruptive automatisms	=	score 4
(e.g. shouting, wandering, undressing)		
Mild automatisms or focal jerking	=	score 2
None	=	score 0
Add 1 point to each column	=	
Total score for each seizure type	=	

* Source: Donoghue MF, Duncan JS and Sander JWAS. (1996). The national hospital seizure severity scale: A further development of the chalfont seizure severity scales. Epilepsia. (37) : 563-571

Conceptual Framework

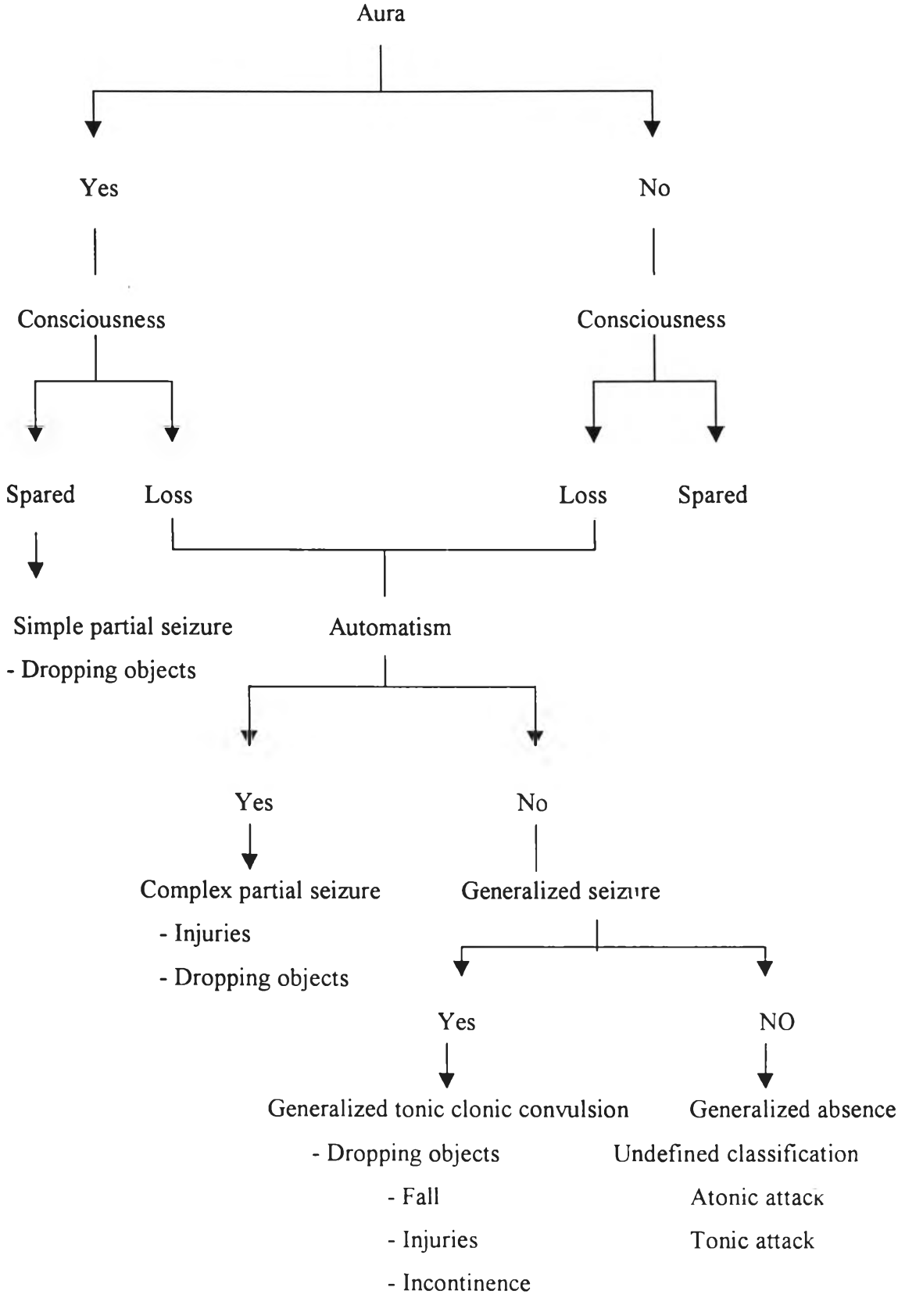
The characteristics of any type of seizure (shown in Figure 6) that can make patients have a difference in severity might be simply classified as ictal and postictal stages.

The ictal stage includes aura (warning), loss of consciousness or automatism that might be present or not. As examples, the patient who has aura before seizure attack might be able to protect him/herself from injuries and the patient with simple partial seizure (spared consciousness) has awareness and can control other parts of body except the affected limb(s). In both he and she can save his or her body which is different from those with seizure with loss of consciousness. Automatism as well, mild automatism might be less severe by witness's perception than seriously

disruptive automatism. Patients with seizure attack during sleep might have less severity than during awake because of fewer injuries occur.

The postictal stage includes the symptoms after ictal stage and time to regain consciousness. Patients might have headache, confusion, fatigue, and feel sleepy after the seizure disappears. Patients who have these postictal symptoms and spend more time for regaining consciousness might have more severity in their feelings than those not having these features.

Figure 6:. Characteristics of Seizure



Objective

- To evaluate the reliability of the CS in terms of interobserver stability

Instrumental Design

The seizure attack will be a variable whether the consideration is the type of seizure and severity of seizure. Standardized protocols for diagnosis and classification of The International League Against Epilepsy are used to identify the type of seizure. The Chalfont Seizure Severity Scale is used to quantify the severity of seizure. The purpose of this scale is to evaluate or measure the effect of treatment. Because there is no gold standard of severity of seizure, this scale was developed with four steps. First, open interviews of 50 patients with epilepsy and their close relatives were conducted to determine what features of their seizures caused disruption and disturbance and whether these were mild, moderate or severe. This enquiry led to a list of 11 factors that were most commonly perceived to be important. Second, these factors were combined into a scale, and each factor was assigned a weighted score, the initial weightings being guided by the results of the open interviews. Third, to adjust the weightings of the individual factors and to assess the content and face validity of the scale, several methods were used:

1. A battery of eight different examples of seizures was drawn up and a seizure severity score obtained for each. The relative severity scores of the different seizure types were then compared with the opinion of a panel comprising the

investigators and six medical and nursing professionals with particular expertise in the management of epilepsy.

2. The scale was piloted on 24 patients with epilepsy. The relative seizure severity scores were compared with the opinions of trained nursing and care staff who had witnessed all the seizures considered. This population encompassed a broad range of patients ranging from those with no neurological or psychological deficit to individuals who had associated mental and physical handicaps.

3. The scale was applied to inpatients and outpatients who had two or more different types of seizure and they and their relatives were asked to compare the perceived relative severity of their seizures, with the severity scores obtained by the scale.

As a result of these processes, the factor weightings were adjusted and the above steps repeated. The final (seventh) version of the scale was then validated using the above steps, and tested in 37 patients who each had more than one seizure type, including patients with simple partial, complex partial, GTC, tonic, atonic, absence and myoclonic seizures. Fourth, inter-rater and test-retest reliability of the final version of the scale were assessed. For inter-rater reliability assessments, patients and a reliable witness were interviewed by the two observers, independently, at the same clinic attendance. Retest reliability was assessed after an interval of two to three weeks and each patient was interviewed with the same witness as previously. Patients were excluded from the test-retest assessment if the

patient or witness subjectively thought that there had been a material change in seizure severity in the intervening period. Statistical analysis of validity and reliability was by the method of Bland and Altman.

The questionnaire is composed of 12 items. The first one asks for the type of seizure that is an open-ended question. The answer is a categorical nominal data. There are 7 alternatives that interviewers can match with the patients' symptoms by using The Standardized protocols for diagnosis and classification of The International League Against Epilepsy and more than one alternative can be chosen. The second to the ninth items ask for whether or not loss of awareness, warning, drop/spill a held object, fall to ground, injury, incontinent, automatism, and convulsion. The tenth and the eleventh items ask for the duration of seizure and the time to return to normal from onset. The last item is the total score calculated by summation of all scores. The numerical discrete score is assigned from the original one. The third to the eighth questionnaire items, the score will be fractionated by the percent of occurrences such as no score if that factor does not occur or quarter score if occurs in up to 25% of occurrences. The questionnaire is shown in Table X.

*Table X: The Chalfont Seizure Severity Scale Questionnaire

Classification of seizure type 1=Absence 2= Simple partial 3=Complex partial
4= GTC 5= Tonic, atonic
6= Myoclonic 7=other.....

1 Loss of awareness	no = 0	yes = 1
2 Warning (if loss of awareness)	no = 1	yes = 0
3 Drop/spill a held object	no = 0	yes = 1 or 2 or 3 or 4
4 Fall to ground	no = 0	yes = 1 or 2 or 3 or 4
5 Injury (tonque biting, bruising, laceration)	no = 0	yes = 5 or 10 or 15 or 20
6.Incontinent (urine ± feces)	no = 0	yes = 2 or 4 or 6 or 8
7. Automatism	no	= 0
	mild (chew, swallow, fiddle)	= 1 or 2 or 3 or 4
	severe (shout, undress, run, hit)	= 3 or 6 or 9 or 12
8 Convulsion(clonic jerk)	no = 0	yes = 3 or 6 or 9 or 12
9 Duration of seizure	<10 sec. = 0	10 sec.-1 min. = 1
	1-10 min. = 4	> 10 min. = 16
10 Time to return to normal from onset	< 1 min. = 0	1-10 min. = 5
	10-30 min. = 20	30-60 min. = 30
	1-3 hrs = 50	> 3 hrs. = 100

Score .....

If epileptic event (eg. brief aura) with total score = 0 , then add 1 .....

Divide score by 2 if only in sleep .....

total .....

(items 3- 8 if there is 25% occurrence = 1/4 score, 50% = half score, 75% = 3/4 score)

* Source: Duncan JS and Sander JWAS. (1991). The chalfont seizure severity scale. Journal of Neurology, Neurosurgery, and Psychiatry. (54) : 873-876.

Validity Test of the Instrument

1. Content validity:

Three neurological experts have evaluated the items of the questionnaire. Their opinions are that the questions were suitable for the epileptic patients and covered all of the important content.

2. Criterion validity:

This validity cannot be evaluated because of no gold standard for prediction the severity of seizure.

Reliability Test of the Instrument

1. Test of internal consistency:

Because of unequal weighting score of each item, Cronbach's alpha, the test of internal consistency with the non-dichotomous score, cannot be used to evaluate. Trying to transform the non-dichotomous score to dichotomous score is done by denoting the occurrence of bad event with zero and the absence of bad event with 1. The duration of seizure equal or less than 1 minute is denoted by 1, otherwise this will be zero. The time to return to normal from onset equal or less than 30 minutes is denoted by 1, otherwise this will be zero. After transformation, Kuder-Richardson will be applied to test the internal consistency.

2. Test of stability:

The inter-observer consistency method is used to assess the stability. There are two observers who performed measurement of this scale at the nearly same time. For nominal data, Kappa is used to assess the stability and Intraclass Correlation is used for continuous data that is calculated by SPSS program.

Pre-Test Plan

Both observers spent time studying the original paper and standardized protocol. After that, the words and scoring the questionnaire were clarified for the two observers so as to get the same meaning on the basis of the original meaning. The permission for testing the instrument with epileptic patients at The King Chulalongkorn Memorial Hospital epilepsy clinic was granted by The Head of Division of Neurology, Faculty of Medicine, Chulalongkorn University and the doctor responsible for the epilepsy clinic.

Method of Using the Questionnaire

Structured interview is used

Data Gathering

At the epilepsy clinic, consecutive epileptic patients gave informed consent and answered the questions. Two observers asked patients independently at nearly

the same time. Patients who have had no seizure attack during last 3 months would be excluded. The original data is in Table XI and Table XII.

Table XI: The Data from Observer 1

ID	Type	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Total
1	4	1	1	0	0	0	0	0	12	1	5	20
2	4	1	0	0	0	0	0	0	12	4	20	37
3	4	1	0	0	0	20	0	0	0	4	30	55
4	4	1	0	0	4	20	0	0	6	16	20	67
5	4	1	1	0	4	20	0	0	12	4	5	47
6	3	1	0	0	1	0	2	1	0	1	5	11
7	4	1	0	0	1	0	0	4	0	4	5	15
8	4	1	1	0	4	0	0	0	12	4	5	27
9	4	1	0	2	3	10	0	0	0	4	5	25
10	4	1	1	0	0	0	0	0	12	4	30	48
11	4	1	0	1	4	0	2	0	3	4	0	15
12	4	1	0	0	0	0	0	0	0	4	50	55

Table XII: The Data from Observer 2

ID	Type	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Total
1	4	1	1	4	0	0	0	0	12	4	5	27
2	4	1	0	0	0	0	0	0	12	1	20	34
3	4	1	0	0	0	20	0	0	0	4	30	55
4	4	1	0	0	4	20	0	0	12	4	100	141
5	4	1	1	0	4	20	0	0	12	4	50	92
6	3	1	0	0	0	0	2	3	0	1	5	12
7	4	1	0	0	0	5	0	0	0	4	5	15
8	4	1	1	0	2	0	0	0	12	4	30	50
9	4	1	1	0	2	5	0	0	0	4	5	18
10	4	1	1	0	0	0	0	0	12	4	30	48
11	4	1	0	0	4	0	0	0	0	4	5	14
12	4	1	0	0	0	0	0	0	0	4	30	35

Statistical Test

In this study two observers performed measurement of variables, therefore, the stability of reliability was tested by using the SPSS program.

Kappa is used to test inter-observer stability of type of seizure. The result is equal to 1. The Intraclass Correlation is used to test inter-observer stability of the first item to the total score. The results are shown in Table XIII.

Table XIII: The Inter-Observer Stability Test

Item	People Mean Square	Observers Mean Square	Residuals Mean Square	F test	Significant	Intraclass Correlation or R.C
Item 1	No standard deviation					
Item 2	0.46	0.04	0.04	1.00	0.33	0.84
Item 3	0.76	0.04	0.95	0.04	0.83	-0.08
Item 4	6.40	1.04	0.22	4.66	0.053	0.93
Item 5	155.30	0.00	2.27	0.00	1.00	1.00
Item 6	0.77	0.16	0.16	1.00	0.33	0.65
Item 7	1.21	0.16	0.89	0.18	0.67	0.15
Item 8	70.73	0.37	2.01	0.18	0.67	0.94
Item 9	9.00	6.00	6.81	0.88	0.36	0.13
Item 10	629.82	759.37	361.64	2.09	0.17	0.27
Total	1455.37	590.04	334.49	1.76	0.21	0.62

For testing the internal consistency of this questionnaire, the data is transformed to dichotomous scale shown in Table XIV:

Table XIV: Transform Data to Dichotomous Scale

ID	Item											
	1	2	3	4	5	6	7	8	9	10	X_t	X_t^2
1	.00	.00	.00	1.00	1.00	1.00	1.00	.00	.00	1.00	5	25
2	.00	1.00	1.00	1.00	1.00	1.00	1.00	.00	1.00	1.00	8	64
3	.00	1.00	1.00	1.00	.00	1.00	1.00	1.00	.00	.00	6	36
4	.00	1.00	1.00	.00	.00	1.00	1.00	.00	.00	.00	4	16
5	.00	.00	1.00	.00	.00	1.00	1.00	.00	.00	.00	3	9
6	.00	1.00	1.00	1.00	1.00	.00	.00	1.00	1.00	1.00	7	49
7	.00	1.00	1.00	1.00	.00	1.00	1.00	1.00	.00	1.00	7	49
8	.00	.00	1.00	.00	1.00	1.00	1.00	.00	.00	.00	4	16
9	.00	.00	1.00	.00	.00	1.00	1.00	1.00	.00	1.00	5	25
10	.00	.00	1.00	1.00	1.00	1.00	1.00	.00	.00	.00	5	25
11	.00	1.00	1.00	.00	1.00	1.00	1.00	1.00	1.00	1.00	8	64
12	.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.00	.00	7	49
Sum	0	7	11	7	7	11	11	6	3	6	$\sum X_t = 69$	
P	0	.58	.92	.58	.58	.92	.92	.5	.25	.5	$\sum X_t^2 = 437$	
Q	1	.42	.08	.42	.42	.08	.08	.5	.75	.5	$S_t^2 = 4.34$	
PQ	0	.24	.07	.24	.24	.07	.07	.25	.18	.25	$\sum pq = 1.61$	

The formula of Kuder-Richardson is

$$r = \frac{n}{n-1}\left\{1 - \frac{\sum p_i q_i}{S_t^2}\right\}$$

$$S_t^2 = \frac{n\sum X_t^2 - (\sum X_t)^2}{n(n-1)}$$

$$r = 0.7$$

Interpretation

1. For testing internal consistency, it cannot be applied by Cronbach Alpha because the weighting score of each item is not equal to other items. When we try to assess with Kuder-Richardson by transforming the data to dichotomous scale, the correlation coefficient is 0.7. It is not quite good.

2. For testing stability, Kappa's value of type of seizure is equal to 1. It means there is no inter-observer variation using the Standardized protocols for diagnosis and classification of The International League Against Epilepsy.

3. When each item score is compared between two observers, the Intraclass Correlation of the second, fourth, fifth, sixth, and eighth item is more than 0.5 but the third, seventh, ninth and tenth item asking about droping or spilling a held object, automatism, duration and time to return to normal is very low. The reason might be:

3.1. From examinees, patients could not perceive or understand some questions themselves, therefore they might guess it. It is necessary to have witness with good memory to get a reliable answer that is the solution of these problems. The other reason is the patients' mental problem contributes to inconsistency of answers.

3.2. From examiners, if the examiner uses only the structured interview, the answer might be incorrect because of patient's ignorance. The examiner might ask the detail of each seizure attack (clarification probe) and might get more data. However, if the patient has multiple seizures there might be memory problem. The other problem is the seizure attack during sleep. It is difficult to assess the duration of seizure and time to return to normal. The onset might be known if someone sleeps with them but the actual time to return to normal might blend with the sleeping time. The time assignment of this questionnaire has no gap such as 10 seconds to 1 minute and 1 minute to 10 minutes and 10 minutes to 30 minutes. If patient stated 1 minute or 10 minutes, what score does the examiner put on the questionnaire.

4. The Intraclass Correlation of total score is 0.56. This questionnaire might apply to the patients but needs to pay attention to the factors that influence reliability.

Discussion

The severity of seizure, a measurement of epilepsy outcome which might be more important than number of seizure as mentioned above (the details on pages 88-89), composes of two main components. One is patient ownself can perceive such as asking whether or not about warning sign before seizure occurrence and about loss of awareness during the seizure occurrence, and the other is a witness can only observe from the patient such as asking whether or not about dropping or spilling a

held object, falling to ground, automatism, and convulsion. Hence, It has to be measured precisely by asking both patient with a witness in order to get more accuracy and reliability. However, the accuracy and reliability might have an inevitable limitation because the patient with mental retardation cannot detail to the attending doctor or seizure occurs during sleep or there is no witness found as well as the number of seizure occurrences might not be perceived by the patient and be observed by a witness. And if it is feasible because of patient's forgetfullness, the patient should be asked to record immediately the numbers and features of seizure after they happened which can increase more accuracy and reliability as well.

REFERENCES

Duncan JS and Sander JWAS. (1991). "The chalfont seizure severity scale". Journal of Neurology, Neurosurgery, and Psychiatry. (54) : 873-876.

Baker GA,... [et al.]. (1998). "ILAE commission report; commission on outcome measurement in epilepsy 1994-1997 Final Report". Epilepsia. (39) : 213-231.

Donoghue MF, Duncan JS and Sander JWAS. (1996). "The national hospital seizure severity scale: A further development of the chalfont seizure severity scales. Epilepsia. (37) : 563-571.