ORIGINAL

片麻痺機能テストの標準化に関する研究 (1)

—— 研究方法 ——

佐久間 昭^{*1} 上 田 敏^{*2} 間 得 之^{*3}
 長谷川 恒雄^{*4} 楠 正^{*5} 木 船 義 久^{*5}

(受付: 1976 年 4 月 9 日)

Standardization of an Objective Test of Motor Recovery in Hemiplegia (I)

Akira SAKUMA^{*1} Satoshi UEDA^{*2} Tokuyuki HAZAMA^{*3}
 Tsuneo HASEGAWA^{*4} Tadashi KUSUNOKI^{*5} and Yoshihisa KIFUNE^{*5}

^{*1} Dept. of Clin. Pharmacol., M. R. I., Tokyo Med. and Dent. Univ.

^{*2} Central Rehabil. Service, Univ. of Tokyo Hosp.

^{*3} Nakaizu Spa Hosp.

^{*4} Izu Nirayama Rehabil. Hosp.

^{*5} Stat. Res. Sec., Development Dept., Takeda Chemical Ind., Ltd.

A multi-institutional cooperative study had been planned in order to standardize the conventional test for motor recovery in hemiplegia. Furthermore, the study group intended to establish a sensitive test which is essential in evaluating a given therapy in hemiplegia. For promoting the study as sure as possible, a fundamental plan was first prepared in which the whole study was divided into sequential Steps. The study started in 1971 and finished in 1975, and included 10 hospitals.

In this report are described the preparatory and methodological considerations (Step 1, 2). After Brunnstrom's Test which takes account of the regularity in recovery process, an initial tentative test was set up that gives 11 grades instead of the original 6 stages. The test is referred to as the 10-Grade-Test in this series of reports since it makes 10 intervals of recovery for both the upper and lower limbs.

^{*1} 東京医科歯科大学・難研・臨床薬理学部門

^{*2} 東京大学医学部付属病院・リハビリテーション部

^{*3} 中伊豆温泉病院・内科

^{*4} 伊豆韮山温泉病院・内科

^{*5} 武田薬品工業株式会社・開発部・試験解析室

The test is constructed by a battery of subtests (test items); 20 for the upper and 23 for the lower extremity. The procedure and the categorical judgement for each subtest were defined, and the criteria for each Grade as well as for each Stage corresponding to that of Brunnstrom's were described. Some definitions are to be altered along with the progress of study.

The methodological considerations cover the principles in design, analysis and assessment for the reliability and for the validity of the test and subtests, and the procedure of quantification of the scale of recovery process.

Key Words : Hemiplegia, Brunnstrom's Test, Reliability of test, Validity of test, Scaling

各種の脳障害によってひきおこされる片麻痺に関するリハビリテーション医学の進展に伴い、治療、訓練などの成果を吟味することを含めて、片麻痺患者を回復過程の上に位置づけるテストについての関心が高まってきた。すでに片麻痺機能テストは、いくつか報告されているが、たとえばBrunnstromのテストはわが国においても広く用いられているものの1つである。

Brunnstromのテスト¹⁾は、片麻痺機能の回復過程が一応の規則性に沿って流れているという彼女の洞察の上に組立てられている²⁾。初期の完全麻痺の状態にあつては、筋緊張も腱反射もほとんど消失した真の弛緩期が認められる。ついで低緊張一反射亢進の時期を経て対側性連合反応を認めるようになる。この時期には随意運動はまだ不可能であるが、健側の強い筋収縮が患側の筋活動を誘発する。

その後、随意運動が出現しても、半ば不随意的な要素をまじえた、原始的運動統合の時期と考えられ、制約された型にはまった共同運動パターンにとどまる。連合反応、姿勢反射が共同運動パターンを種々の型で修飾する。またこのころには、筋緊張は多少とも増強に向かうことが多い。

ついで、共同運動パターンから、部分的に独立分離した個別の運動が次第に可能になる。この過程では、まず単純な中枢性促進の回復がおり、ついで、これに加えて選択的な中枢性抑制が可能に

なるものと解釈できる。個々の運動の分離が完成するとともに、協調運動が正常化する時期に向かう。

およそ、こうした規則性を前提にして、片麻痺患者をStage IからStage VIまでの回復過程という尺度上に分類するものがBrunnstromのテストである。Stage IからIIIまでが共同運動パターの完成、Stage IVからVIまでが分離動作の完成の時期にあたる。

このテストは原理的にも経験的にも、ほぼ適切なものとして広く受け入れられ、実用されているが、なお、若干の問題点をもっている。たとえば、上肢についての検証はほぼ十分と思われるが、下肢については不十分である。また、Stageやサブテストの具体的な規定に不明確な点が残されており、さらに、詳細な治療成果の吟味という立場からみると、Stage IIIを中心に、より細分化した評価の基準が望まれる。

われわれは、Brunnstromのテストの原理にもとづき、多施設協同の研究として、片麻痺機能テストの標準化とStageの細分化を試みた。この研究は1971年5月に開始されたが、本来は武田薬工企画開発本部によるCDPコリンの薬効評価の準備作業として計画されたものである。1975年6月の時点において、テストの標準化と一応の検証を終り、所期の目的にはほぼかなった片麻痺機能テストの基準案を得たので、ここに報告する。ここで

は、全般的な計画、データ解析を通じての尺度構成、テストの構成といった面を重視し、リハビリテーション医学³⁾、ならびに薬効評価⁴⁾の観点からの検討は、別報にゆずることとする。なお伊豆菰山温泉病院、竹田総合病院、中伊豆温泉病院、兵庫リハビリテーションセンター、洞爺協会病院、九州労災病院、厚生年金玉造病院、甲州中央病院、鹿教湯温泉病院、東京大学リハビリテーション部の10施設が今回の臨床研究に参加した。

基本計画

片麻痺機能テストの標準化という複雑な規模の大きな研究課題をとりあげるにあたり、協同研究者間に共通の目標の理解を徹底させるように努めた。研究が論理的に進められるよう、研究の出発にあたり、参加者によって検討を行い、Table 1のようなStepに分けて、進行を図ることにした。実際には同時的に扱ったものもあり、その暦日を付記してある。

Step 1は、従来の報告、臨床経験をもとにした考察であり、Step 2はテスト構成の試案の作製段階であった。Step 3以降が臨床レベルでの観察とデータの吟味を行う段階で、Step 3では、ほぼStage IIIにあたる患者について、評価者間の判定を比較し、一致性を吟味した。Step 4, 5は同時的に実施したが、各施設での評価者内、評価者間の判定の一致性、日間変動の検討を行った。こうした信頼性についての吟味に加え、テストが回復過程を正しく把握しているかどうかの妥当性を吟味した。Step 6, 7では、さらに妥当性の吟味を行うとともに、治療、訓練による回復過程での動き幅について、より検証的な吟味を行った。

Table 1 基本計画

Step 1: 過去のデータの収集と分析	1971. 5~1971. 10
Step 2: 評価方法の検討→第1回計画検討会	1971. 9~1972. 2
Step 3: 評価の一致性の検討(Stage III)(中伊豆温泉病院での検討会)	1972. 2~1972. 4
Step 4: 測定値の一致性の検討(全Stage)(各施設でのデータ収集)	1972. 5~1973. 6
Step 5: 片麻痺機能テスト試案の妥当性の検討	
Step 6: 一定期間内の片麻痺回復の動き幅の検討	1973. 6~1975. 6
Step 7: CDPコリン投与による影響の検討(二重盲検比較対照試験)	

尺度構成という立場からは、Step 1, 2が準備段階で、Step 3, 4, 5が信頼性と妥当性をふまえての基本的な尺度作りの段階で、Step 6, 7が、より検証的な感度の吟味を中心とした段階といえるだろう。今回の報告も、この線に沿って、(1)研究方法、(2)基礎的吟味、(3)定量的吟味と検証の順に述べることにする。

Grade 試案とサブテスト

ここでいう片麻痺機能テストとは、片麻痺患者を回復過程の上に位置づけるテスト全体を指し、このテストは多数の具体的なサブテストの組で構成されている。このサブテストは心理学でいうテスト項目(アイテム)にあたる。

具体的なサブテストを念頭において、Brunnstromに準じたStageの細分化の可能性と意義とを検討した結果、Table 2のようなGrade試案を作製し研究の出発点とした。実際には0から10まで11個のきざみがあるが、動きが10個ということから、これを10段階Grade試案と呼ぶことにする。

随意収縮の出現から共同運動パタンの完成にかけて、上肢では屈筋、下肢では伸筋の運動が先行すると思われたが、ここでは上下肢とも共通に扱い、屈伸の2本立てとした。BrunnstromのStage規定には不明確な点が残されているため、今回の研究では、上下肢とも、Table 2に示す規定によってStageを定めている。

この10段階Grade試案と平行して、サブテストとその対象になる筋を検討し、具体的なサブテストの手技、結果の判定基準、記録方式をまとめ、上下肢につき、片麻痺機能テストの評価表を作製した。研究の進行に応じて検討を行い、若干の改訂を加えたが、ほとんどが観察、判定の技術的な問題に関係し、テストの本質的な変更はなかった。

上肢で20、下肢で23のサブテストを用意した。サブテストの判定は、不可能0と可能1の2枝択1の場合と、可能の程度を細分し1, 2, 3, ……などに分類する多枝択1の場合とがある。サブテストの中の選択枝0, 1, 2, ……などをここで

Table 2 10段階 Grade 試案

Brunnstromに 準ずる Stage	Grade
I	0 筋収縮なし
II	1 対側性連合反応による筋収縮
II	2 随意収縮の出現。但し関節の動きなし
III	3 共同運動パターン、屈曲・伸張の一方のみ出現
III	4 共同運動パターン、屈曲・伸張の両方出現。但し不十分
III	5 共同運動パターン、屈曲・伸張の両方、十分に出現
IV	6 分離動作(Brunnstromに準ずるStage IVのテスト動作)の1つが十分
IV	7 分離動作(Brunnstromに準ずるStage IVのテスト動作)の2つが十分
V	8 分離動作(Brunnstromに準ずるStage Vのテスト動作)の2つが十分
V	9 分離動作(Brunnstromに準ずるStage Vのテスト動作)の3つが十分
VI	10 分離動作(Brunnstromに準ずるStage Vのテスト動作)の3つが十分で、 スピードテストが健側の1/2~1/3以上 注

注) スピードテスト: 患者と健側についてテスト動作が5秒間に何回できるかを測定する。Step4以降では逆に5回もしくは10回の動作に何秒要したかを測定することにしたため、表3の表現とここの表現とは異なる。

は判定のカテゴリーと呼ぶことにし、どの程度に可能ならば十分に可能とみなすかの十分規定とあわせて Table 3 に示した。Table 2 の Grade 試案ならびに Table 3 の十分規定に基づいて対象の Grade が定められる。

サブテストは最終的には上下肢とも11個にしばれそうであるという見通しを得たが、しばらくは、この当初のサブテストの組について記す。

サブテストの配列は、ほぼ易しいものから難しいものへと並べたが、実施体位の便宜性を加味した。また共同運動パターンについては、屈曲を先行すると伸張が抑制されがちであるといった経験から、伸張パターンを先行させるようにした。

上肢のサブテストは Brunnstrom のテスト¹⁾とほとんど同じであるが、実施体位を変更している。下肢のサブテストは Brunnstrom のテストのもの、ならびに上田、福屋の報告⁵⁾に準じたものを含んでいる。

なお、片麻痺機能テストの実施にあたっては、拘縮によって正確な評価が妨げられるおそれがあるため、他動的関節可動域の制約をおいた。その他、日常生活動作(ADL)テストに準じた歩行中心の下肢基本動作テストなども実施している。

Table 3 片麻痺機能テストの概要

(上肢)		判定カテゴリー			
テストの動作	テストNo.	0	1	2	3
伸屈連合反応	1a	なし	ATNR 併用で あり	単独で あり	
屈曲連合反応	2a				
大胸筋の収縮	1b-1				
上腕三頭筋の収縮	1b-2				
僧帽筋の収縮	2b-1				
上腕二頭筋の収縮	2b-2	不可能	耳〜乳頭 〜臍	乳頭〜臍 〜乳頭	臍〜伸屈位 乳頭〜耳
三角筋の収縮	2b-3				
伸屈共同運動	1c				
屈曲共同運動	2c				
手を背中へ	3				
腕を前方水平位に挙上	4	不可能	可能 背中より5cm以内	0〜25°	30〜55°
肘を体側につける	5-1				
肘の屈内	5-2				
肘の屈外	5-3				
腕を前方上方に挙上	6				
肘伸展位で屈内	7-1	可能	中間位(〜55°)まで可 可能	90〜125°	130°〜
肘伸展位で屈外	7-2				
肘伸展位で腕を横水平位に	8				
テスト8の動作	9				
手を肩から頭上に挙上	10				

(下肢)

レイミストの連合反応(股外転)	1-1	なし	あり		
レイミストの連合反応(股内転)	1-2				
股内転筋群の収縮	2a-1	なし	ATNR 併用で あり	単独で あり	
大腿四頭筋の収縮	2a-2				
下腿三頭筋の収縮	2a-3				
股外転筋群の収縮	3a-1				
股屈筋群の収縮	3a-2				
ハムストリング筋の収縮	3a-3	不可能	90〜50°	45〜25°	20°以下
膝の伸張	2b				
膝の屈曲	3b				
背臥位	4				
股関節の外転	5				
定関節の背屈	6	不可能	5〜15°	20°〜	
股関節の屈曲	7				
膝関節の伸張	8				
膝関節の屈曲	9				
定関節の背屈(膝屈曲位)	10				
定関節の背屈(膝伸展位)	11	可能	5〜25°	30〜45°	50°〜
股関節の内転	12				
定関節の背屈	13				
膝関節の屈曲	14				
股関節の外転	15				
テスト11の動作	16	健側の1.5 倍以上	1.5倍以下	5°以上	20°〜
テスト13の動作	17				
	18				

注) 太線以上が十分

下肢テストNo.6の点線はStep6以降の結論による十分

テストが具備すべき条件の検討

片麻痺機能テストを構成する Table 3 の各サブテストの判定結果に応じて、回復過程という1本の尺度上に観察対象を位置づける方法を標準化したいというのが研究目標である。

好ましいテスト、好ましい尺度が備えるべき性質に関しては、古くから心理学の領域で論じられている⁶⁾。われわれもこうした議論を基本において、片麻痺の臨床に特有な問題を念頭におきながら研究を進めた。各サブテストについて、さらにこれらの結果の総括によって与えられる尺度値について、測るべきものが精度よく正確に反映されていないなければならない。

いずれの回復過程にある対象についても、テ

トが与える対象の位置に関して、同じ評価者がくり返しても、異なった評価者がくり返しても、同じ結果が再現されなければならない。回復過程を評価していると思われるサブテストの組を合理的に折半し、同時に、あるいは時をおいてテストを行ったとき、相互の結論がよく一致していることも、信頼におけるテストの条件になる。

一致性が高い、再現性が強いということは、さまざまな回復過程にある対象を互いに弁別し、また同一対象の経時的な動き幅を敏感にとらえることの必須条件になる。さらに、この一致した値が、真値を正しく指していることが要求される。こうした一連の問題が、テストの信頼性にかかわる。

さまざまな回復過程にある多数の対象の、それぞれの真値がわかっているという理想的な情況において、真値の分散を σ_o^2 と示す。これらの対象にテストを実施して得た尺度の観測値と真値とは多少ともずれたものになろう。こうした観測した尺度値の分散を σ_i^2 と示し、真値からのずれにかかわる誤差分散を σ_e^2 と示せば、観測値の性質に若干の条件を設定したときには

$$\sigma_i^2 = \sigma_o^2 + \sigma_e^2$$

と書ける。心理学において、信頼性係数

$$r_{tt} = \frac{\sigma_o^2}{\sigma_i^2} = 1 - \frac{\sigma_e^2}{\sigma_i^2}$$

が定義されているが、観測値全分散に占める対象間分散の割合で、決定係数、関与率に対応するものである⁶⁾。誤差分散が0であれば、信頼性係数は明らかに $r_{tt}=1$ である。

信頼性の高いテストを構成するサブテストの組が得られ、このテストによって確固たる尺度値が与えられるとして、そこで測っているものが、正しく回復過程という医学的内容をとらえているかどうかが問題になる。このテストが、回復過程の全体を覆うかどうか；片麻痺の臨床という観点からみて、十分にうなずける結果を与えるかどうかなどの吟味を必要とする。このようにテストが測っている実質科学的内容にかかわるものが、テストの妥当性の問題である。

この研究においても、サブテストの組が連続的な1次元性をもつ回復過程の上に尺度を張り、一

致性の高い尺度値を与えることのほかに、サブテストの医学的な意味での相互関連性の吟味、間隔的な尺度の観点からの吟味、同一対象の回復の動きを確実に敏感にとらえる性質の付与なども重要な課題としてとりあげた。

今回の片麻痺機能テストの標準化、回復過程の尺度構成といった研究の場合、素材データを手にした後は、その枠内での作業、検討に限られるため、基本計画に述べたように Step の区切りを設け、回顧的に前進するという方針をとった。

すでに確固たる外的基準があり、各対象の位置づけが行われ、簡便法としてのテストを構成するというのであれば話しは変ってくる。今回はサブテストの吟味をしながら尺度を構成し、尺度を吟味しながらサブテストを検討するという試行錯誤的、反復的な接近法によって、テストにおける内部矛盾を除きながら、かつ、片麻痺回復の規則性という仮定に照らしながら研究を進めた。

信頼性の検討方針

ある対象Sを回復過程という尺度上に位置づけるにあたり、現実に生じうる誤差、ゆれうごきの内容は、評価者内、評価者間での評価の不一致による。また一方では、対象Sにテストをくり返したときに疲労、馴れ、真の回復過程での動きなどが絡んでくるため、SがS'に変る可能性を否定できない。こうした現実問題を加味して、評価の一致性を検討する方法を考えてみる。

Fig. 1 には4種類の方法をあげたが、

(A)では同一評価者が、ほとんど同時的に同一対象を評価する。対象の負担、くり返しによる $S \neq S'$ の変化、 x_1 が x_2 に及ぼす影響などを考慮する必

Fig. 1 一致性の検討方法

要がある。

(B)では異なった評価者が、同時的に同一対象を評価する。テストの厳密な同時性、 x_A と x_B 相互の影響などを考慮する必要がある。

(C)では日を改めて、同一の評価者が同一の対象を評価する。(A)と類似の考慮を必要とするが、片麻痺回復の遅さを考えれば(A)より好ましいだろう。

(D)では日を改めて、異なった評価者が同一の対象を評価する。ここでも(B)、(C)と類似の考慮が必要となる。

本研究にあっては理論的な厳密な一致性ないし信頼性の吟味が不可能であるという事情から、改善の策を講じなければならなかった。Step 3では、ほぼ StageIIIにあたる対象について(B)法に準じて評価者間の一致性を吟味した。Step 4, 5では広範囲の Stage にわたる対象について、各施設の事情に応じて(C)、(D)法に準じて一致性ならびに日間変動を吟味した。

評価者内、評価者間の一致性、日間変動の吟味は、各サブテストについて、およびこれらが与える Grade という尺度値について必要になる。その際、サブテストが2枝択1の2つの判定カテゴリーをもち、0, 1の分類データを与えるか、たかだか多枝択1の判定カテゴリーをもち、0, 1, 2, ……などの順序データを与えるにすぎないため、解析上の扱いに若干の問題はあった。

本来の真値が求められないため、ここでは操作的に扱い、すべての対象について、すべての評価が一致したときの完全な一致が100%、また完全な不一致が0%となるような一致率を定めて用いた。

一方では、判定カテゴリーの0, 1, 2, または Grade の尺度値などの生データを用い、やや強引ではあるが分散分析による評価者の比較、日間変動の検討などを行い、一致性に関する考察をおこなうことにした。

一致性の吟味と対象との関係についていえば、あるサブテストが対象の回復の程度により、誰がみても明らかに可能または不可能であるとき、結

果の一致性は当然のことである。そのサブテストについての一致性の吟味は、真実にはそのサブテストについて可能と不可能の中間に位置するような対象が好ましい。このため、個々のサブテストの一致性の吟味に際しては、それぞれのサブテストにふさわしい範囲を設定し、事後的にデータを選択した。

一方、信頼性係数 r_{tt} については、現実には求められない真値というものが関係してくるため、多数のくり返して得られた平均値を操作的に真値と考え、信頼性係数に対応させた表現として決定係数を用いることにした。一致率と決定係数との間には、ほぼ平行的な関係があるものと思われる。

また、直値が広くばらついた対象群に関しては、ある程度の一致性によっても決定係数は大きくなる。逆に、真値がよく揃った対象群に関しては、一致性のわずかなずれによって決定係数は小さくなる。Step 3では対象を StageIII中心にしぼるため、決定係数よりも各サブテストの一致率の吟味が主となり、Step 4では一致率と決定係数とをかみあわせた検討が主となるものと考えられる。

妥当性の検討方針

サブテストが2枝択1的で、可能と不可能、もしくは十分と否、の2つの判定カテゴリーをもつとき、このサブテストが回復過程の尺度にきざむ1点を考えてみる。この1点を境界にして、より初期の、回復の浅い対象と、より後期の、回復の進んだ対象とを区別できるはずである。

ある対象群について、あるサブテストの結果が十分となるものの割合は、そのサブテストに関する、この対象群の合格率、通過率であるが、ここでは十分(者)率と呼ぶことにする。理想的な状況において、このサブテストがきざむ尺度上の境界点よりも初期の対象群では十分率が0%、より後期の対象群では十分率が100%になろう。Fig. 2(A)は、このような理想的なサブテストの特性曲線で、回復過程上のある1点を境にして十分率が0から100%に急変する。

ある指定した1組のサブテストのそれぞれが、

Fig. 2 理想的な特性曲線

Fig. 3 十分率曲線

回復過程の尺度に固有のきざみをもち、Fig. 2(B)のように、ほぼ一樣の間隔で並ぶようであり、しかも固有の点で十分率が0から100に急変しているようならば、理想的なテスト、つまりサブテストの組が与えられたことになる。サブテストについていえば、より初期に対応した易しいものから、より後期に対応した難かしいものへと序列的に並び、本質的には連続性ないし1次元性が成立っている。

一方、現実の場でサブテストの組をつくり、ある対象にテストを実施し、各サブテストが十分か否かに応じて1点か0点を与え、その対象の合計点を計算し、これを操作的な回復過程の指標、尺度値にしてみる。この合計点、心理学でいう素点の順に対象を分類し、多数の対象に関して、それぞれのサブテストの十分率をプロットすると、Fig. 3のようになる。こうした曲線をサブテストの十分率曲線と呼ぶことにする。

理想的なFig. 2(B)のような、境界線で急変する曲線群とは異なって、現実にはFig. 3(A), (B)のような曲線群となるであろう。Fig. 3(A)では、すべてのサブテストについて、合計点が大きくなるにつれて十分率が高くなる単調上昇の傾向があり、比較的急峻な姿になっている。これに対して Fig.

		サブテスト			
		1	2	3	4
反応型	0	0	0	0	0
	1	1	0	0	0
	2	1	1	0	0
	3	1	1	1	0
対象合計点	4	1	1	1	1
	サブテスト合計点	4	3	2	1

Fig. 4 項目得点行列

3(B)では、一部に非単調的な中だるみや山型の十分率曲線を示すサブテストも含まれ、Fig. 3(A)に比べて、テストとして好ましくない。

現実のテスト構成にあたり、できるだけ単調に急峻に変化する十分率曲線をもった、切れ味のよいサブテストを選び、十分率50%のレベルについて、ほぼ一樣のきざみを与えるようなサブテストの組をつくり、かつ医学的な観点から重要で、実施面での問題が少ないサブテストを加えることが必要であろう。この方針に沿って、サブテストの取捨、修正、並べかえをくり返し、より好ましいテストに接近することができる。

上述の Fig. 2(B)の理想的なテストの場合、サブテストの結果の十分か否かを1, 0として、各対象について合計点をつくり、その小さなものから大きなものへと並べかえる。この1や0の並び方を反応型という。一方、サブテストごとに、合計点をつくり、その大きなものから小さなものへと並べかえる。理想的な場合、サブテストは易から難に並ぶ。

もっとも単純な、理想的な場合を考え、それぞれの反応型に1例ずつが該当し、サブテストが易から難に一樣に並ぶとすれば Fig. 4のような、項目得点行列ができる⁷⁾。

この行列の要素は、サブテストの結果0,1であるが、理想的な場合、対角線を境界にして上方に0、下方に1が分離して並び、シンプレックス型と呼ばれる。この項目得点行列の扱いは、多枝の判定カテゴリー0, 1, 2, ……をもったサブテストを含む場合にも拡張できる。現実のデータは、

こうした理想的、決定論的なパターンを示さず、行と列を入れかえても、シンプレックス型とは異なり、1の領域に0が、また0の領域に1が互いに混入したものになる可能性が強い。こうした乱れを誤反応という。一般には行和（反応型）と列和（サブテスト合計点）に応じて並べかえたうえでの項目得点行列の誤反応を問題にする。

いま n 例の対象につき、 t 個のサブテスト ($j=1, \dots, t$) を実施し、上述のような項目得点行列をつくる。それぞれの反応型の例数を考慮して、各サブテストについて誤反応を示した例数 e_j を数えあげたうえ、Guttman の再現性指数と呼ばれる値

$$\text{Rep} = 1 - \frac{\sum e_j}{tn}$$

を計算する。心理学の領域では、この値が 0.85～0.9 をこえるならば、そして誤反応が特定のサブテストに集中していないようならば、これらのサブテストが原則的には 1 次元性の尺度を構成するものとみてよいとされる。

各サブテストごとに $\text{Rep}_j = 1 - e_j/n$ をみれば、1 次元性の観点から問題を含んだサブテストが指摘できるであろう。しかしこの Rep_j は 0.5 より小さくならず、その値はサブテストの本来の難易にも影響をうける。このため、2 枝択 1 のサブテストについては、Jackson のプラス・パーセンテージ・レシオを西里が修正した PPR* という指標⁷⁾も併用した。PPR* は 1～0 の値をとり、0.7 をこえるようならば、1 次元性を想定してよいとされている。

乱れのある項目得点行列については、反応型を中心に横にみると、たとえば 1010 のように、第 2 番目に易しいサブテストが不可能のまま第 3 番目のサブテストが可能という、跳びこしの現象が認められる。

ところで、サブテストによって、片麻痺回復過程という尺度上で対象を測っているつもりであるが、そこで測られている実質は医学的にみてどのようなものであろうか、この種の吟味には因子分析の手法が用いられるが、テストの背景に潜在的な因子なるものを想定し、ある因子がいずれのサ

ブテストを介して顕在化されるのか、各サブテストが、それぞれどのような潜在因子にかかわりあうのかを吟味するわけである。

われわれは、広義の因子分析の 1 つと考えられる主成分分析を用い、サブテスト間の相関行列について解析した。これは項目得点行列を、各サブテストごとに平均値 0、分散 1 に基準化し、対象のばらつきを、より低い次元の世界に映しかえ、ばらつきを代表するような新しい軸、主軸を求めるという固有値問題になる。もとのばらつきが主軸に集約された大きさの割合を寄与率というが、少数の主軸によって 1 に近い累積寄与率をみれば、十分な集約が行われたとみてよい。

この主軸に対する各サブテストのかかわりの程度を示す重み係数（因子負荷量にあたる）を求め、その結果をさらに見易くするためにバリマックス回転を行い、各主軸（因子）に対するサブテストの重み係数を求めることにした。この各主軸に対する重み係数のパターンによってサブテスト相互の関係を洗えば、サブテストが引き出している潜在的な性質をうかがうことができるであろうと考えた。

相関行列を用いると、サブテストの十分率、つまり難易をとりはらっていることになるから、サブテストの難易の情報は薄れるであろう。そこで、必要に応じて、項目得点行列の平均値を引くことなしに求めた積和行列を用い、変則主成分分析（特異分解ないし N-法）⁸⁾を併用したが、これによって、サブテストの難易という性質が、いずれかの主軸に反映されることを期待した。

間隔尺度の検討方針

サブテストの組が 1 次元的な尺度をきざみ、そこで測っているものが医学的に妥当であるなら、一応の尺度作りは成功したことになる。一般に、こうしたテストが対象に与える尺度値は、たかだか順序をもった名義的なものである。

この尺度のきざみに、何らかの間隔的な意味をもった数量が与えられるなら、間隔尺度として、より定量的な対象の位置の検討が可能になる。こ

の種の手法を数量化、尺度化などと呼んでいる。

サブテスト（テスト項目）の結果が0,1のみで示される2枝の判定カテゴリーをもつ場合、また0, 1, 2, ……の順序あるいは質的分類のカテゴリーをもつ場合、これら2枝、多枝のサブテストが混在する場合などがある。これらに対していくつもの数量化の手法があるが、それらの名称は必ずしも統一されていない。2枝の判定カテゴリーをもつサブテストにおいては、数量化に際してはカテゴリーとサブテストが同義的に扱われる。

類似したサブテストないしはカテゴリーには類似した数量を与え、またテストに対して類似した反応を示す対象には類似した数量を与えるという、内的整合性が最大になるように数量化が行われる。異なった反応型をもつ対象には互いに離れた位置が与えられるので、対象の識別という要求が満たされる。ここでいう間隔尺度とは、この種の基準にしたがったものを指している。

いま、2枝択1のサブテストの組について、 n 例の対象が p 種類の反応型（0, 1の並び方のパターン）のいずれかに属したとしよう。そのとき、理想的な項目得点行列は Fig. 4 のシンプレックス型になるが、実際には、反応型としてみてもサブテストとしてみても、0, 1の乱れがあるであろう。

これらの反応型に $x_1, \dots, x_i, \dots, x_p$ の数量を与える。一方、2枝のサブテストが t 個あったとして、これらには $y_1, \dots, y_j, \dots, y_t$ の数量を与える。ここで x_i, y_j といった数量を与えるに際して

$$(1) \eta^2_{xx} = \frac{x \text{ の反応型間の分散}}{x \text{ の全分散}}$$

$$(2) \eta^2_{yy} = \frac{y \text{ のサブテスト間の分散}}{y \text{ の全分散}}$$

$$(3) \eta^2_{xy} = \{x \text{ (反応型) と } y \text{ (サブテスト) との相関係数}^2$$

のそれぞれを最大にしてみる。これら3種類の基準によって x_i, y_j を定めることは解析上は固有値問題になり、いずれの場合も同じ数量 x_i, y_j を与え、 $\eta^2_{xx} = \eta^2_{yy} = \eta^2_{xy} = \eta^2$ となる。3つの η^2 の値

は形式上は異なるが、いずれも決定係数、関与率にはかならない。ここでは η^2_{xx}, η^2_{yy} を相関比と呼んでおくが、 η_{xx}, η_{yy} を相関比と呼ぶこともある。

こうして求めた x_i, y_j は間隔尺度として扱えるが、必要ならば符号を逆にして反応型を x_i の大きさに並べると、回復過程に関して何らかの間隔的な意味をもたせて、初期から後期へと位置づけができる。同様に y_j についてサブテストを並べると、間隔的に意味をもたせて易と難とが区別できる。

サブテストが多枝で、その判定カテゴリーに数量を与える場合も、ほぼ同様の原理で扱うことができる。ここではサブテストのカテゴリーに数量を与える方法を Guttman の最適尺度法⁷⁾、サブテストに数量を与える方法を林の数量化第3類⁹⁾と呼んでおく。

各サブテストないしカテゴリーに与えられた数量を比べると、医学的に矛盾した逆順になっていることもある。この際にはサブテストないしカテゴリーの除去、あるいは併合を行って再検討を試みる。2つのサブテストなりカテゴリーが接近しているときには、いずれかを除いても実際上は問題がなく、テストの簡潔化を図ることができる。

反応型、サブテストに合理的な数量を与えることができたならば、医学的な観点からの検討を加え、いわば準間隔尺度といったものを構成し、ある対象 i に一定期間の治療、訓練を行い、回復過程上の尺度値の動き幅 d_i について、より定量的な検討が可能になる。また、この動き幅 d_i はその対象の年齢、性別、罹病期間、治療や訓練の種類、初回の尺度値などによって、どのような影響をうけるかをうかがうことができる。

対象 i における動き幅 d_i を基準変量、応答変量とし、これを左右すると思われる年齢、性別など q 個の説明変数 $w_{i1}, \dots, w_{iq}$ を考えたうえ、重回帰（一般線形回帰）モデルにしたがい最小2乗法の原理によって、

$$\hat{d}_i = \hat{\beta}_0 + \hat{\beta}_1 w_{i1} + \dots + \hat{\beta}_q w_{iq}$$

の線形式をあてはめ、それぞれの説明変数の意義

を吟味する。

この際に、動き幅に曲線的な影響を与えたり、頭打ちの現象を生じる可能性のある説明変数については、 $w_{ik}' = (w_{ik} - \bar{w})^2$ といった加工を行い、質的な分類をダミー変数として扱うことにする。

小 括

片麻痺機能の回復過程に患者を位置づける Brunnstrom のテストには、実用面に関して若干の問題が残されている。このテストが与える 6 段階の Stage をより細分化し、標準化することを目標として多施設協同の研究を企図した。

片麻痺回復の規則性を前提におき、11 個のきざみをもつ Grade 試案を作成したが、これには動きとしての数が 10 個あるので、10 段階 Grade 試案と呼ぶことにして、上下肢それぞれに 20 個、23 個の具体的なサブテスト（テスト項目）を用意して片麻痺機能テストを構成するという方針を研究の出発点とした。

協同研究者の間に、研究の方針と意義、研究の手順を徹底させることに努力し、研究過程を Step として区切り、基本計画として整理したうえ、回顧的に前進することにした。

テストの信頼性と妥当性の吟味をくり返し、テストが与える回復過程の尺度に間隔的な意味をもたせ、より定量的な感度のよい、医学的な内容を

反映し、かつ実用に耐えうる尺度を構成するための検討方法を討議した。この第 1 報の内容は、基本計画でいう準備段階の Step 1, 2 にあたる。

文 献

- 1) Brunnstrom, S.: Movement Therapy in Hemiplegia, A Neurophysiological Approach, Harper & Row, New York, (1970). 佐久間穰爾, 松村秩 (訳): 片麻痺の運動療法. 医歯薬出版, 東京, (1974).
- 2) 上田敏: 成人片麻痺の評価. 総合リハビリテーション, 3: 717-725 (1975).
- 3) 上田敏, 福屋靖子, 間得之, 長谷川恒雄, 佐久間昭: 片麻痺機能テストの標準化に関する研究. 総合リハビリテーション, 投稿予定.
- 4) 間得之, 長谷川恒雄, 上田敏, 佐久間昭: 脳卒中後遺症に対する CDP-choline の多施設二重盲検法による薬効評価. 臨床と研究, 投稿予定.
- 5) 上田敏: 目でみるリハビリテーション医学, 東京大学出版会, 東京, (1971).
- 6) Guilford, J. P.: Psychometric Methods. McGraw-Hill, New York, (1954). 秋重義治 (監訳): 精神測定法 (初版). 培風館, 東京, (1959).
- 7) 西里静彦: 応用心理尺度構成法. 誠信書房, 東京, (1975).
- 8) Okamoto, M.: Four Techniques of Principal Component Analysis. 日本統計学会誌, 2: 63-69 (1972).
- 9) 林知己夫, 樋口伊佐夫, 駒沢勉: 情報処理と統計数理. 産業図書, 東京, (1970).