

精神分裂病に対する日本版 DAS の 評価者間信頼性の検討について

東京医科大学精神医学教室 (指導: 清水宗夫主任教授)

本 郷 誠 司

【要旨】 能力障害評価面接基準 (DAS) は、精神医学的な機能障害および能力低下の評価基準の作成、その自然経過、機能レベルの予測因子の同定を目標とした「精神医学的能力障害の評価と軽減に関する WHO の協同研究」のために開発され、1991 年に日本語に翻訳された。

DAS は 5 つの区分、61 の項目で構成された構造化面接のひとつで、各国で追試が行われているが、全項目での追試の報告はない。そこで、2 カ所の施設、32 名の精神分裂病患者に対して全項目に渡って追試した。結果に対する信頼性の統計学的な検定には、 κ 統計量を用いた。

κ 統計量による一致率の検討では、DAS 全項目 61 項目中の 46 項目 (全体の 75.4%) で高い評価者間一致度を得た。評価者間一致度の低い項目は、第 3 区分に集中 (15 項目中 12 項目: 80.0%) したが、それは「基準分散」によると思われる。その他の項目 (15 項目中 3 項目: 20.0%) は、「観察分散」によると思われる。

以上の結果から第 3 区分のアンカーポイントの再設定や 6 段階評価への統一などの内容の再検討・修正が早期に必要と思われる。しかしながら、それらを加味しながらの DAS の使用は、臨床応用上有益であると思われた。

はじめに

精神疾患患者の精神医学的評価尺度は今日までに数多くの研究や発表がなされているが、その多くが試行錯誤の過程を現在も行っている¹⁻⁷⁾。例えば身体疾患の場合、正常値の数値が病状改善や治療効果の指標となる数量的指標の一つとして存在する。他方精神疾患の場合、身体疾患の正常値のような数量的指標が一般に乏しく、とくに精神分裂病のような重度の精神障害の総合的な判定には困難が多いためである。しかし、精神障害者の社会復帰が叫ばれている現況ではより客観的で適正な病状改善・治療効果・社会適応性・予後などの把握・評価・判定が要求される。

1991 年に丸山らにより翻訳された能力障害評価面接基準 (Disability Assessment Schedule 以下 DAS と略す)⁸⁾ は、精神医学的な機能障害および能力低下の評価基準の作成、その自然経過、機能レベルの予測因子の同定を目標とした「精神医学的能力

障害の評価と軽減に関する WHO の協同研究」⁹⁾ のために開発された。DAS 以前には、臨床的あるいは精神療法的場面における対人行動を評価する多くの尺度や評価基準が使用されたが、主観的な相違から評価者間にかなりの食い違いがあり、また精神障害者の社会復帰に関しての指標を調べ検討する基準や、精神分裂病のような重度の精神疾患を伴う患者の社会機能全般を評価するために開発された評価方法が無かった。DAS は精神障害者の社会行動の評価法間におけるギャップを埋めるために注意深く設計され、同時にできる限り客観視、数量的に扱うように配慮・作成されたが、日本ではそのような研究調査は無かった。

社会的な機能の障害を、一疾病の症状と捉えるかあるいは能力障害の範疇と解釈するのかという明確な区別を下すのは困難であるが、区別すること自体に身体的なりハビリテーション¹⁰⁾や精神科的リハビリテーション¹¹⁾に対して有益であることが示されている。WHO による医学的な機能障害、能力障害

1995 年 1 月 26 日受付, 1995 年 6 月 6 日受理

キーワード: 評価者間信頼性, 能力障害評価面接基準, 精神医学的能力障害, 精神分裂病, κ 統計量。

(別刷請求先: 〒160 東京都新宿区西新宿 6-7-1 東京医科大学精神医学教室 本郷誠司)

および社会的不利についての国際分類¹²⁾によれば、「機能障害 (impairment)」とは「心理学的、生理学的、あるいは解剖学的な、構造または機能における、何らかの欠如または異常 (any loss or abnormality of psychological, physiological, or anatomical structure or function)」とされている。そして「能力障害 (disability)」は「人間にとって正常とみなされるやり方、または範囲内で、活動を達成しうる能力の何らかの制限および欠如である (any restriction or loss of ability to perform an activity in the manner or within the range considered normal for a human being)」と定義されており、「機能障害」に伴う社会生活技能や職業技能の障害を意味している。DAS における「能力障害」は「その生活環境において個人が普通に期待されている特定の社会的な役割を遂行する能力の欠如あるいは制限」として、更に一段と専門的に定義されている⁹⁾。つまり精神医学的能力障害とは、「精神疾患を有する者の、その生活環境において健常者に期待されている特定の社会的な役割を遂行する能力の欠如あるいは制限」であると解釈される。

DAS の内容は大きく 5 つの区分 (表 1) からなり、更に区分毎に下位区分や小区分が存在する。第 1 区分は「全般的行動」で、4 つの下位区分を有し、第 2 区分は「社会的な役割の遂行」で 10 個の下位区分からできている。第 3 区分は「入院患者」のみについて、4 つの下位区分 (病棟内の行動、看護婦の意見、患者の仕事、外界との接触) が存在し、その各々に更に 12, 9, 4, 4 の計 29 個の小区分からなっている。第 4 区分は「修正要因」で、4 つの下位区分 (特別な資質、特別に不利な点、家庭環境、外での支持) と、

その各々に 5, 3, 6, 3 の計 17 個の小区分が存在する。第 5 区分は「包括的評価」で、1 項目となっている。

DAS の実地研究は 7 カ国 (ブルガリア、西ドイツ、オランダ、スーダン、スイス、トルコ、ユーゴスラビア) の研究センターで特定の診断基準を満たす 520 名の患者を対象に行われ、評価者間信頼性、妥当性が検討された。そこから得られた結果は評価者間の一致度や異なる社会文化、国際的な比較研究にも利用できることが示された⁹⁾。DAS は、その後イタリアで実施¹³⁾されたり日本でも氏原¹⁴⁾により実施されている。しかしながら、研究報告は DAS の総ての項目に対してではなく、特に日本においては DAS の入院患者に対して評価する部分である「第 3 区分」が得られていない。その「第 3 区分」の「4 つの下位区分」 (病棟内の行動、看護婦の意見、患者の仕事、外界との接触) の内で、前二者は看護職員から得られた情報による評価である。DAS の重要な目的の一つに、医師のみでなく、コメディカルによる患者評価をも可能にする、ということがある。そこで、今回その「第 3 区分」の入院患者に対する評価や看護職員による評価という部分を含め、日本版 DAS を用いて、総ての項目における評価者間信頼性の検討を試みた。

対象と方法

対象は、評価者が所属する 2 カ所の施設 (東京医大病院: 以下 TMC, 富士リハビリ病院: 以下 FR と略す) の 32 名の精神分裂病および分裂病様障害と診断された患者を対象に行われた。その内訳は表 2 のごとくであった。患者および家族には研究の主旨を説明し協力を仰ぎ了解を得た。

実際の方法は、2 名の評価者が第 3 区分の 21 項目を除く 8 項目と、他の 4 区分 32 項目、計 40 項目について、同席面接法で 1 名の患者または家族の面接に臨んだ。一人が主に面接を行ないながら同時に評価し、もう一人は面接のやり取りを観察しながら評価を下した。また、観察側が、面接者の不足と思われる質問を追加した。面接の際には DAS に付属する「面接の為の手引き」を、内容に留意しつつ加筆変更した手引きを作成・利用しながら、手順に従って施行し、評価は互いを意識せず考慮せずに独立して下した。第 3 区分の看護職員から得た情報による評価 (3. 1 の 12 項目および 3. 2 の 9 項目) の 21 項

表 1 DAS の構成内容

区 分	項目数
第 1 区分: 全般的行動	4
第 2 区分: 社会的な役割	10
第 3 区分: 入院患者	病棟内の行動 12 看護婦の意見 9 患者の仕事 4 外界との接触 4
第 4 区分: 修正要因	特別な資質 5 特別に不利な点 3 家庭環境 6 外での支持 3
第 5 区分: 包括的評価 評価の要約	1

表 2 対象患者内訳

施設別詳細					平均年齢±標準偏差	
TMC	9 例	男	4 例	入院	4 例	25.00±10.95
				外来	0 例	
		女	5 例	入院	3 例	34.00±17.32
				外来	2 例	26.50±7.78
FR	23 例	男	17 例	入院	13 例	31.92±11.67
				外来	4 例	25.00±7.39
		女	6 例	入院	0 例	
				外来	6 例	35.67±8.62

目については、一定の 2 名の看護職員評価者にアンケート形式の書類(表 3)を渡し、互いに相談などせずに記入してもらった。結果の信頼性の検定には、Cohen の κ (カッパ) 統計量¹⁵⁾を用いた。評価者間信頼性の計算方法は多く存在^{16)~19)}するが、今回の研究では

- 1) アンカーポイントが設定されている。
- 2) データが数値化されている(カテゴリーカルなデータでない)。
- 3) 評価者が 2 名である。

という前提で実施した。これらの条件下で一致率の信頼性を計算する方法には、最近の多くの研究で使用されている Cohen の κ 統計量¹⁵⁾が最適である。その定義は「『見かけの一致率』から『偶然の一致率』を差し引いたもの」で、つまり偶然による一致の補正を行なうという点で、%一致率などの他のものより優れている。そして κ 統計量が 0.6 以上あれば通常一致率が良いという指標であると考えられている¹⁵⁾。

κ 統計量を数式で示すと次のようになる。

$$\kappa = (p_o - p_c) / (1 - p_c)$$

p_o = 見かけの一致率 (%一致率)

p_c = 偶然の一致率

2 名の評価者の評価が完全に一致、換言すれば偶然の一致が全く無ければ κ 統計量は 1 になり、 κ 統計量が 0 であれば評価は偶然の一致でしかない。

κ 統計量の計算過程で %一致率が求められるが、%一致率では偶然による評価の一致も含まれている可能性がある。故に数値的には %一致率より κ 統計量は常に低値であり、同じ %一致率でも κ 統計量が異なる値となることがある。結果には %一致率も参考までにあげた。なお実際の統計処理及び数値演算にはパーソナルコンピュータを使用して、評価

者 2 名のマトリックスを生成して行った。%一致率は、%で表示するために小数点以下 4 桁とし、 κ 統計量はそのままの数値を表示するために、小数点以下 3 桁を有効数字とした。

結 果

項目順に得られた結果が表 4 であるが、 κ 統計量について総括的に見ると、第 1 区分の 4 項目は全て 0.6 以上と評価者間一致率は高かった。同様に第 2 区分の 10 項目中、3 項目は完全に一致し、0.6 未満の評価者間一致率の低い項目は 1 項目であった。第 3 区分の 29 項目中、12 項目では 0.6 未満であった。第 4 区分の 18 項目では、2 項目で 0.6 未満であった。第 5 区分の 1 項目では、 κ 統計量が 0.754 となった。

結果の内では完全に一致した 6 項目、すなわち κ 統計量が 1 である項目と、 κ 統計量が 0.6 未満を示した一致率が低いと思われる 15 項目について以下に述べる。

1. 完全に一致した項目

完全に一致した 6 項目中、3 つは第 2 区分で、2 つが第 3 区分、1 つが第 4 区分であった。

第 2 区分は『社会的な役割の遂行』についての能力障害の評価で、家事への参加、夫婦の役割、親の役割、社会的な接触、職業、緊急時の行動などに対する評価を下している。その内で完全に一致した 3 つの項目は、2. 2 の「夫婦間の役割: 過去一カ月間における配偶者との情緒的な交流 (ありふれたでき事などについて話し合う、愛情や暖かさを示す)」と、2. 3 の「夫婦の役割: 過去一カ月間における配偶者との性的な関係 (過去一カ月間の性交の回数、性的な満足を経験したか、配偶者が性的な満足をしたか)」、及び 2. 4 の「親としての役割: 過去一カ月間

表 3 看護スタッフに対するアンケート

患者_____さんの病棟内の行動について (記入看護スタッフ名_____)

以下の項目について、過去 1 週間の患者さんの行動を考えて 1 つ選択して数字を○で囲んで下さい。

1. 動きの遅さ

- 0 動きの速度は普通である。
- 1 顕著な遅さが長時間見られる。しかし、普通の速さで動けることもある。
- 2 例えば食事をとったり着替えあるいは病棟内を歩く時に、他の患者よりずっと時間がかかるなど、極端な動きの遅さや動作の緩慢さが常に見られる。

2. 活動性の低下

- 0 著明な活動性の低下は見られない。
- 1 顕著な活動性の低下が長時間見られる。しかし、それらが見られない時もある。
- 2 一ヶ所に殆ど動かずに立ったままかあるいは座ったまま居て、催促しても動かすことは非常に難しい。

3. 活動性の増加 (多弁も考慮に入れる)

- 0 著明な活動性の増加や多弁は見られない。
- 1 顕著な活動性の増加や多弁が長時間見られる。しかし、それらが見られないこともある。
- 2 例えば歩く速度が一定でなかったり、興奮して大声で乱暴に喋ったり、歌ったりするように、いつも極端に活動性の増加が認められるか多弁あるいは落ち着かない状態である。

4. 会話

- 0 普通の会話をしていた。
- 1 例えば、質問をしたり欲求を訴えるのに数語で済みます。しかし、他の場合は沈黙している。
- 2 無言、あるいは殆ど喋らない。

5. 社会的引きこもり

- 0 普通に打ち解けて交流できていた。
- 1 周囲の社会や環境から自分自身の内に引きこもっていたり、孤独であったりする。しかし、促されれば他の人と消極的ながら交わって行ける。
- 2 促されても誰とも打ち解けて交わろうとしない。

6. 楽しみに対する興味

- 0 普通に自発的な興味を示す。
- 1 殆ど興味を示さない。しかし、勧められると何とかテレビを観たり、新聞を読んだり、ゲームに参加したりする。
- 2 いかなる事にも興味を示さない。例え勧められても、テレビを観たり、新聞を読んだり、ゲームに参加することはない。

7. 筋違いの或はまとまりのない話し

- 0 そのようなことは、見られなかったし記録されてもない。
- 1 時には筋違いの事を喋ったり、まとまりのないことを言ったりすることがある。しかし、これらのことは毎日起こる訳ではない。
- 2 筋違いのこと、あるいはまとまりのないことを言ったりすることがしばしばある (1日に1回以上)。

8. 奇妙な姿勢あるいはわざとらしさ

- 0 そのような行動は見られないし、記録もされていない。
- 1 一見奇妙な、あるいは不快な姿勢をとったり、あるいはわざとらしい行動をするが、毎日という訳ではない。
- 2 毎日奇妙で不快な姿勢をとったり、あるいはわざとらしい行動をする。

9. 脅威を感じさせる粗暴なふるまい

- 0 そのような行動は見られないし、記録もされていない。
- 1 態度は脅かすようなそぶりを示したり、言葉は口汚いが、誰も傷つけはしない。
- 2 誰かを傷つけ、何かを破損したりする (例えば衣類、窓、机、椅子、食器など)。

10. 閉じこもったり、ベットにもぐり込んでしまう傾向

- 0 かなり離床していることが多い。
- 1 ベットや布団の中にいたり、ベットや布団に戻ってしまう傾向がある。しかし、離床を促せばそれに応じられる。
- 2 離床を促されても、執拗にベットや布団の中にもぐり込んだままで、離床してもすぐに戻ってしまう。

11. 個人的外見

- 0 促されたり、監督されることなしに、外見をかなり小奇麗に保ていられる。
- 1 髪を剃ったり（男性の場合）、化粧をしたり（女性の場合）、衣類を着たり、入浴することができる。しかし、注意や監督が必要で、外見がむさ苦しくなりやすい。
- 2 1 週間に多くてもやっと 1 回、髪を剃ったり、化粧をしたり、入浴したり、着替えたりする程度。

12. 食事の際の行動

- 0 食事の際の行動は普通である。
- 1 食事の介助の必要はなかったが、食べ散らかすなどの為に幾分監督が必要である。
- 2 この 1 週間に少なくとも 1 回は、食事の介助が必要だった。

13. 患者さんは、もし可能なら、病院内で洗濯や掃除などの家事的な仕事が出来ますか？

- 0 はい。 1 いいえ。 2 該当無し。

14. 患者さんは、マッチやライターを持たす事が可能ですか？

- 0 はい。 1 いいえ。 2 該当無し。

15. 患者さんは、週末に自由に外出や外泊がさせられますか？

- 0 はい。 1 いいえ。 2 該当無し。

16. 患者さんは、異性の患者さんと外出させられますか？

- 0 はい。 1 いいえ。 2 該当無し。

17. 患者さんは、ハサミやカミソリの刃を持たせられますか？

- 0 はい。 1 いいえ。 2 該当無し。

18. 患者さんは、自分でお金を管理できますか？

- 0 はい。 1 いいえ。 2 該当無し。

19. 患者さんは、院外作業などの仕事に従事できますか？

- 0 はい。 1 いいえ。 2 該当無し。

20. もし仕事が見つかり、適当な住む所があれば、患者さんは即刻退院できますか？

- 0 はい。 1 いいえ。 2 該当無し。

21. 患者さんは、開放病棟や開放的な部屋で適応してやって行けそうですか？

- 0 はい。 1 いいえ。 2 該当無し。

における子供への関心と養育（両親の地位に相応しく、子供の養育の役割を引き受けたり、実施している。子供への関心）」であった。

第 3 区分は『入院患者に対する評価』で、過去 1 カ月間に少なくとも 2 週間入院している患者について評価を下す。病棟内の行動、看護婦の意見、病棟内外での活動、外界との接触などに対する評価である。その内でも完全に一致した 2 項目は、病棟内の行動に関する 3. 1. 12 の「食事の際の行動」と、外界との接触に関しての 3. 4. 2 の「過去 3 カ月間に自宅へ外泊したか？」であった。

第 4 区分は、『修正要因』の検討で、特別な資質、特別に不利な点、家庭環境、外部での支持などに対する評価である。その内で完全に一致した項目は、家庭環境に関しての 4. 3. 2 で、「過去 1 カ月間内の

平均的な一週間で、前記の中心人物（4. 3. 1: 患者にとって過去 6 カ月間における家庭での中心人物はだれか？）とどの位面と向かって接しているか？何時間？」であった。

2. χ 統計量が 0.6 未満を示した項目

χ 統計量が 0.6 未満を示した項目は、全部で 15 項目である。これらを χ 統計量の高い順から低い順へ向かって以下に述べる。まず第 1 に、2. 1 の「過去一カ月間の家事への関与（一緒に食事をしたり、家事に関するような雑用をしたり、外出したり、家庭における普通の活動に関わっているか。家計上の決断に参加しているか。主婦の場合は、主婦が通常行っている家事を念頭に置く。）」である。2. 2 や 2. 3、2. 4 が完全に一致したのに対して、2. 1 は一致率が低いと言える。

表 4 DAS 各項目の結果

		%一致率	κ
1. 全般的行動			
[1.1]	過去 1 ヶ月間の患者の自己管理	81.25%	0.715
[1.2]	過去 1 ヶ月間の活動性の低下	81.25%	0.733
[1.3]	緩慢さ	78.13%	0.685
[1.4]	過去 1 ヶ月間における社会的な引きこもり	84.38%	0.773
2. 社会的な役割の遂行		%一致率	κ
[2.1]	過去 1 ヶ月間の家事への関心	68.75%	0.599
[2.2]	夫婦間の役割：過去 1 ヶ月間の情緒的交流	100.00%	1.000
[2.3]	夫婦の役割：過去 1 ヶ月間の性的な関係	100.00%	1.000
[2.4]	親としての役割	100.00%	1.000
[2.5]	性的な役割	78.13%	0.716
[2.6]	社会的接触：人間関係の摩擦	81.25%	0.739
[2.7]	職業上の役割：過去 1 ヶ月間の仕事の達成	84.38%	0.809
[2.8]	職業上の役割：過去 1 ヶ月間の仕事への関心	81.25%	0.728
[2.9]	過去 1 ヶ月間の興味と情報	75.00%	0.674
[2.10]	緊急時の行動	78.13%	0.712
3. 入院患者			
3. 1. 病棟内の行動		%一致率	κ
[3.1.1]	動きの遅さ	85.00%	0.653
[3.1.2]	活動性の低下	75.00%	0.531
[3.1.3]	過動（多弁も考慮に入れる）	95.00%	0.810
[3.1.4]	会話	80.00%	0.619
[3.1.5]	社会的な引きこもり	75.00%	0.556
[3.1.6]	楽しみに対する興味	80.00%	0.579
[3.1.7]	勘違いのあるいはまとまりのない話	80.00%	0.630
[3.1.8]	奇妙な姿勢あるいはわざとらしさ	80.00%	0.588
[3.1.9]	脅威を感じさせる粗暴なふるまい	95.00%	0.773
[3.1.10]	閉じこもったりベッドにもぐり込む傾向	80.00%	0.596
[3.1.11]	個人的外見	75.00%	0.519
[3.1.12]	食事の際の行動	100.00%	1.000
3. 2. 看護婦の意見		%一致率	κ
[3.2.1]	軽作業が可能か？	90.00%	0.744
[3.2.2]	火の始末が可能か？	80.00%	0.583
[3.2.3]	外泊が可能か？	70.00%	0.537
[3.2.4]	外出が可能か？	75.00%	0.632
[3.2.5]	刃物の管理が可能か？	80.00%	0.579
[3.2.6]	金銭の管理が可能か？	90.00%	0.758
[3.2.7]	院外作業が可能か？	85.00%	0.760
[3.2.8]	退院が可能か？	75.00%	0.578
[3.2.9]	解放病棟での適応が可能か？	95.00%	0.903
3. 3. 患者の仕事		%一致率	κ
[3.3.1]	病棟あるいは家庭内における家事に関する活動	76.92%	0.655
[3.3.2]	病棟外での作業療法	84.62%	0.792
[3.3.3]	作業療法あるいは産業療法的活動	76.92%	0.690
[3.3.4]	他のソーシャル・セラピー（芸術活動など）	76.92%	0.695
3. 4. 外界との接触		%一致率	κ
[3.4.1]	過去 3 ヶ月間に面会があったか？	92.31%	0.768

[3.4.2]	過去 3 ヶ月間に自宅に外泊したか?	100.00%	1.000
[3.4.3]	安全の面での管理が必要か?	61.54%	0.409
[3.4.4]	強力な管理が必要か? 理由は?	61.54%	0.532
4. 修正要因			
4. 1.	特別な資質	%一致率	κ
[4.1.1]	特別な領域についての関心	84.38%	0.626
[4.1.2]	趣味についての関心	84.38%	0.688
[4.1.3]	特定の領域についての平均以上の能力や資格	96.88%	0.937
[4.1.4]	好影響を及ぼしている環境や境遇	78.13%	0.536
[4.1.5]	安定した対人関係	87.50%	0.789
4. 2.	特別に不利な点	%一致率	κ
[4.2.1]	不利な環境	93.75%	0.846
[4.2.2]	身体的・神経学的障害	93.75%	0.817
[4.2.3]	不利な特徴 (性格障害, 外見, 習癖)	93.75%	0.765
4. 3.	家庭環境	%一致率	κ
[4.3.1]	中心人物	93.75%	0.897
[4.3.2]	中心人物との接触時間	100.00%	1.000
[4.3.3]	中心人物による情緒的関わりの程度	75.00%	0.646
[4.3.4]	中心人物による期待, 要求, 圧力	81.25%	0.683
[4.3.5]	中心人物による拒否の程度	84.38%	0.745
[4.3.6]	プライバシーの有無	87.50%	0.779
4. 4.	外部での支持	%一致率	κ
[4.4.1]	外部からの手助けや援助	87.50%	0.762
[4.4.2]	実際に受けた援助の有無	96.88%	0.896
[4.4.3]	就労の可能性	62.50%	0.493
5.	包括的評価	81.25%	0.754

以下第 3 区分の 8 項目が続く。3. 1. 10 の「閉じこもったりベッドにもぐりこんでしまう傾向」、3. 1. 8 の「奇妙な姿勢あるいはわざとらしさ」、3. 2. 2 の「患者にマッチを持たせることができるか?」、3. 1. 6 の「楽しみに対する興味」、3. 2. 5 の「ハサミやカミソリの刃を持たせられるか?」、3. 2. 8 の「仕事が見つかり適当な住むところがあれば、即刻退院できるか?」、3. 1. 5 の「社会的な引きこもり」、3. 2. 3 の「週末に自由に親戚や友人を訪問させられるか?」である。次は 4. 1. 4 の「患者の機能的レベルに好影響を及ぼしている何か特徴的な環境あるいは最近の境遇があったか?」であった。

更に第 3 区分の 3 項目が続き、3. 4. 4 の「もし強力な管理が必要なら、その理由を述べよ」、3. 1. 2 の「活動性の低下」、3. 1. 11 の「個人的外見」の順で低かった。

そして次に 4. 4. 3 の「最近の社会文化的な状況に関する全般的な評価に基づいて、患者と同じ年齢、性、教育程度があり、精神疾患の既往のないものが、現在、賃金労働を得ることがどのくらい難しいと面接者は判断するか?」であった。

最後に 3. 4. 3 の「安全の面で管理が必要か?」であった。

考 察

本研究結果を総括的に見ると、 κ 統計量が 0.6 以上の評価者間一致率の高いものは全 61 項目中の 46 項目 (全体の 75.4%) となり非常に良好な一致率であると言える。

1. 完全に一致した 6 項目について

完全に一致した項目の 2. 2, 2. 3, 2. 4 の 3 つについては、「夫婦の関係」や「親としての役割」の評

価の部分である。これらの項目は、「患者からの協力的な情報を必要とする」という但し書きが「使用の手引き」に記されているように、非常に個人的な情報である。特に夫婦間の情緒的な関わりや性生活などは、日本においては、性習慣や羞恥心などからごく表面的な画一的な情報しか得られない可能性がある。同時に被験者 32 名中、既婚者 3 名 (9.38%) とごく少数であったことが、完全一致という結果に結びついたと思われる。

完全に一致した項目の 3. 1. 12 は「食事の際の行動」についてで、その回答の選択肢は

0. 食事の際の行動は正常である。

1. 食事のマナーが良くないので、幾分監督が必要である。

2. この 1 週間に少なくとも 1 回は、食事の介助が必要であった。

の 3 通りである。この項目の被験者 20 名中、19 名が正常と評価され、1 名が幾分監督が必要であると評価された。食事の際の看護職員の注意と対応は、主に介助を要する患者に注がれていることが多い。それ故に評価者である看護職員の患者に対する印象や記憶は比較的明快であるために、完全一致の結果になったと思われる。

3. 4. 2 の「自宅への外泊」は、患者の外泊は医師の指示で許可が出されるが、一般に看護記録の必要性などから看護職員が出院、帰院などを管理すると思われる。故に看護職員の評価に関する情報は、一元的で確実なものとなり、完全に一致したと思われる。

4. 3. 2 は、「中心人物 (大半の時間を患者と一緒に過ごす人物) とどのくらい接触 (中心人物が実際そこにいる) していたか?」、その時間を尋ねている。2 桁の数値を結果として記入するのだが、質問するときに評価者から「1 日 4 時間くらいで、週平均 28 時間くらいですか?」などと時間の例を具体的に挙げざるを得ず、自ずと同一の結果が得られたために完全に一致したと思われる。

2. 一致率の特に低かった項目について

0.6 未満の項目 (15 項目) については一致率が低いと思われるが、その中の 80.0% にあたる部分 (12 項目) が、第 3 区分の評価にあたる。一般に評価の不一致の要因は 5 つ存在し、「被験者分散」「状況分散」「情報分散」「基準分散」「観察分散」と言われている¹⁹⁾。

「被験者分散」とは、患者の症状変化などにより生ずる。例えば急激な病状の悪化などにより、同一患者であっても少し前の状況とは一変していることがある。これは、実際に配慮すべき変化であったりするので、対処は無いと思った方がよい。

「状況分散」とは、被験者である患者の状態は変わりはないのだが、評価者の時間的、空間的な差違により生ずる。例えば同じ患者を午前と午後、外来と病棟、1 対 1 と 1 対複数で評価するなどである。これは同席面接で防ぐことが可能である。今回の研究では、2 名の評価者が同席面接を行っているので排除できていると思われる。

「情報分散」とは、被験者である患者に対する情報の違いにより生ずる。例えば患者自身に質問をした場合や親族・知人に尋ねた場合の情報量は自ずと違っており、評価も変化してしまう。これは、標準化された質問と回答の選択肢により防ぐことができる。DAS では、この辺りを配慮してある。

「基準分散」とは、同一の情報に対してそれを統合する方法や判断するための基準が評価者によって異なっているために生ずる。例えば、「午後 11 時から午前 6 時まで眠れる」と答えた患者の言葉を、7 時間は良く眠れていると解釈したり、早朝覚醒傾向があるなどと解釈したり、評価者によって違った解釈をしてしまうことがある。これは、判断するための基準を明文化し、たとえば先の睡眠時間の解釈も、患者や患者を取り巻く環境の普段の様子を基にして 2 時間以上早く起きれば早朝覚醒傾向有りとする、3 時間以上睡眠時間が短縮されれば睡眠障害有りとするなどの解釈の程度を限定することで防ぐことができる。

「観察分散」とは、「基準分散」に似ているが、同一の情報に対してその重症度や閾値の判定が異なることにより生ずる。

結果の内特に一致率の低い部分が第 3 区分に集中したのは、他の区分が 6 段階評価が中心 (0 から 5 の 6 段階で、情報が得られない場合と核当しない場合も含めると 8 段階となる) であるのに対して、第 3 区分では異なっているための『基準分散』によると思われる。第 3 区分の第 1 小区分 12 項目が 3 段階評価、第 3 区分の第 2 小区分 9 項目が 2 値評価、第 3 区分の第 3 小区分では 3. 3. 1 が 4 段階評価、3. 3. 2 と 3. 3. 3 が 6 段階評価、3. 3. 4 が 5 段階評価で、第 3 区分の第 4 小区分では 3. 4. 1 と 3. 4. 2 が 3 段

階評価, 3. 4. 3 が 5 段階評価, 3. 4. 4 が 8 段階評価となっている。なお, 第 3 区分の第 2 小区分, 第 3 小区分ならびに第 4 小区分では, 「核当無し」という評価項目が以上の段階評価以外に追記されている。このように第 3 区分では複数の評価段階が存在し, かつ第 1 区分や第 2 区分と評価段階が異なっているために一致率が低くなったと思われる。このような基準分散に対する対処法は, 「基準を明文化する」ことである。第 3 区分の評価段階に対して「基準を明文化する」ということは, 他の区分と同様に 6 段階評価に修正したり, 6 段階評価できるような質問内容にする, ということである。基本的に第 1 区分や第 2 区分では, 「能力障害無し」「軽い能力障害」「明らかな能力障害」「重度の能力障害」「極度の能力障害」「最高度の能力障害」の 6 段階となっているが, 第 3 区分では, 必ずしもそうっていない。その理由の一つには, 入院患者の病棟内での行動の乱れを看護職員からの情報や患者の能力に対して看護職員の意見を聞かねばならず, 特に今回の研究では看護職員の意見を得るためにそれらを表 3 のアンケート形式とし, 回答の選択肢はオリジナルのままで行った, という点が挙げられる。このアンケートの 21 項目中の約半数近い 10 項目において一致率が低い。3. 1. 1 から 3. 1. 12 の 12 項目に関しては, 3 段階評価であるために, χ^2 統計量においては 3 段階評価での 1 段階の違いと 6 段階評価での 1 段階の違いよりは, 「1 段階の違い」に「重み」を生じたためであると思われる。また, 3. 2. 1 から 3. 2. 9 の 9 項目は, 基本的には 2 値評価であるが, 結果として 3 段階評価と同様の「重み」が生じてしまい, 「基準分散」となったと考えられる。それらの点で, 一致率を上げる一つの方法として第 1 区分や第 2 区分のような 6 段階評価に修正する必要がある。しかしながら, アンケートと同席面接とは自ずと評価の方法に違いが生じていることは否定できず, 看護職員と評価する医師とが同席で面接する方法などでも改善は期待できるので, 追って実施してみたい。なお第 3 区分では, 3. 4. 3 と 3. 4. 4 が一致率が低い。これらは外界との接触に関しての項目で, 3. 4. 3 は「安全の面で管理が必要かどうか」, 3. 4. 4 は「もし強力な管理が必要なら, その理由を述べよ」となっている。3. 4. 3 が実は一致率が一番低値であったが, これは評価者による主観度の違いから生じた「観察分散」によるものと思われ, 同席面接法などでは避け

られないものである。3. 4. 4 は 3. 4. 3 と密接な関係があるので, 同様に低値となったのであろう。

一致率の低かった項目の第 3 区分を除く残りの 20.0% (3 項目) は, 2. 1 と 4. 1. 4 および 4. 4. 3 である。

2. 1 は「患者の家事への関心」であるが, 評価に当たっては家族の一員としての機能と義務が考慮されなければならない (DAS 第 2 部使用のための手引きより)。役割特性 (たとえば, 妻, 父, 息子などとして) があるとしても, それは家族の間で異なり, すでに確立した形式に基づいている (DAS 第 2 部使用のための手引きより)。それらを考慮するための患者の家族背景などは, 家族と面談をしている主治医とそうでない医師とでは主観度が違い, 容易に「観察分散」を生じさせてしまうと思われる。これは同席面接法などでは避けられないものと思われる。

同様に 4. 1. 4 の「患者の機能的レベルに好影響を及ぼしている何か特徴的な環境あるいは境遇」や 4. 4. 3 の「最近の社会文化的な状況に関する全般的な評価に基づいて, 患者と同じ年齢, 性, 教育程度があり, 精神疾患の既往のないものが, 現在, 賃金労働を得ることがどのくらい難しいか」という項目においても, 評価者間の「観察分散」から一致率が低値になったと思われる。

3. 他の DAS 研究との結果比較について

他の研究者による DAS 研究との結果比較については, 比較可能な第 1 区分と第 2 区分の 14 項目における χ^2 統計量の検討では, ほぼ同様の満足の行く結果が得られたと思われる (表 5)。

4. χ^2 統計量のパラドックスについて

本研究の全体として評価は, χ^2 統計量で検討する限りでは良好な信頼性が得られたと思われるが, 1993 年に友田²⁰⁾によって紹介された Feinstein らの χ^2 統計量に関する 2 つのパラドックス²¹⁾²²⁾から, 単一の総括的指標としての χ^2 だけで評価者間信頼性を検討するのではなく, 分離指標の導入の検討が必要と思われ, また同時に第 3 区分のアンカーポイントの設定や 6 段階評価への統一, 評価方法などの検討及び修正が, DAS の臨床応用には急務であると思われる。それらが利用可能となれば, 医師のみでなく, 精神障害者を援護する多くのコメディカルの客観的な患者評価が可能となり, 精神障害者の社会復帰の前進という点で, 非常に有益であろうと思われる。

表 5 他の研究結果との κ 統計量の比較

DAS 項目		本郷ら	氏原ら*	Ardoïn ら**	Groningen***	Mannheim***	Nottingham***
患者の自己管理	[1.1]	0.72	0.76	0.61	0.74	1.00	1.00
活動性の低下	[1.2]	0.73	0.84	0.77	0.86	0.89	0.90
緩慢さ	[1.3]	0.69	0.58	0.79	—	—	—
社会的な引きこもり	[1.4]	0.77	0.56	0.58	0.85	1.00	1.00
家事への関与	[2.1]	0.60	0.46	0.82	0.73	0.70	0.79
夫婦間の役割：情緒的交流	[2.2]	1.00	0.74	0.78	—	—	—
夫婦間の役割：性的な関係	[2.3]	1.00	0.74	0.78	1.00	1.00	0.24
親としての役割	[2.4]	1.00	1.00	0.76	—	—	—
性的な役割	[2.5]	0.72	0.44	0.68	—	—	—
社会的接触：摩擦	[2.6]	0.74	0.64	0.83	—	—	—
職業上の役割：仕事の達成	[2.7]	0.81	0.71	0.88	0.93	0.91	0.63
職業上の役割：関心	[2.8]	0.73	0.46	0.69	—	—	—
興味と情報	[2.9]	0.67	0.62	0.52	0.73	0.81	0.69
緊急時の行動	[2.10]	0.71	0.63	0.90	—	—	—

(*は文献 14 より, **は文献 13 より, ***は文献 9 より)

5. 本研究全体に関して

今回の研究では、DAS の信頼性の検討のみであったが、一般的に多評価者間による評価尺度の検定においては、その信頼性と妥当性の検討が必要であるとされている。WHO における妥当性の検討は、精神分裂病患者群と対照群の DAS の各項目のプロフィールを比較することにより、両群で明らかな差異が生じることから妥当性があると報告している。また、Schubart らは精神科医とそうでない調査者との 5. 包括的評価の項目における相関係数が 0.79 であったことより DAS の妥当性を主張している。氏原は DSM-III-R の第 V 軸である GAF 尺度を用いて、DAS との併存妥当性を相関係数を用いて検討している。当然本研究でも妥当性の検討が必要ではあるのだが、研究を始めた時点や本研究の革新性などの観点から、妥当性の検定に値する評価が他に見あたらないと判断し、本研究では敢えて信頼性の検定のみとした。しかしながら、今後機会あれば妥当性の検討も加味したい。ちなみに氏原の併存妥当性における GAF 尺度と DAS との相関係数は、5. 包括的評価の項目において -0.75 と高いものであった。

ま と め

1) 日本語版能力障害評価面接基準（日本版 DAS）を、分裂病圏の患者 32 名に対して施行した。

2) κ 統計量による一致率の検討では、DAS 全項目 61 項目中の 46 項目（全体の 75.4%）で高い評価者間一致度を得た。

3) 評価者間一致度の特に低い項目は、第 3 区分に集中（15 項目中 12 項目：80.0%）したが、それは「基準分散」によると思われた。その他の項目（15 項目中 3 項目：20.0%）は、「観察分散」によると思われた。

4) κ 統計量のパラドックスから、評価者間一致度の検討に際しての単一の指標としての κ 統計量の使用には、十分な注意と検討を要すると思われる。

5) 第 3 区分のアンカーポイントの設定や 6 段階評価への統一、評価方法などの検討・修正が必要と思われる。

6) 同一症例の経時的な変化・再評価などの縦断的な研究も今後の課題であると思われる。

7) 本研究では妥当性の検定は行わなかったが、諸家の報告より妥当性は高いと思われる。

8) 上記を加味しつつ DAS を使用することは、臨床のみでなく社会復帰の現場に携わるスタッフや医師以外のコメディカルにとっても有益であると思われる。

稿を終るにあたり、終始御指導、御校閲を賜りました恩師清水宗夫教授、加藤正明名誉教授、並びに協力、助言をくれた同僚丸田敏雅先生、アンケート調査に協力してくれた富士リハビリ病院の看護職員に深謝致します。また、本研究の機会と資料を快く御教示下さった国立精神衛生研究所の丸山晋博士に深謝致します。

なお本研究の主旨は、第 14 回日本社会精神医学会において発表した。

文 献

- 1) 納谷敦夫, 横山淳二, 岡田英明, 浅尾博一: 慢性分裂病患者の病棟における行動評価. 臨床精神医学 **11** (8): 1004~1011, 1982
- 2) 北村俊則, Kahn A., Kumar R.: 慢性精神分裂病の評価尺度. 慶應医学, **59**(3): 385~400, 1982
- 3) 北村俊則, 島 悟: 慢性精神分裂病における陰性症状評価尺度の評価者間信頼度. 慶應医学 **61**(3): 277~283, 1984
- 4) 北村俊則, 島 悟ら: 慢性精神分裂病に対する臨床評価尺度の信頼度. 精神医学 **29**(9): 933~940, 1987
- 5) 尾崎 新, 三宅由子: 慢性分裂病に対する「回復」評価尺度の開発. 精神科治療学 **3**: 735~743, 1988
- 6) 北村俊則, 島 悟ら: 慢性精神分裂病の陽性症状と陰性症状. 精神医学 **31**(2): 131~136, 1989
- 7) 山田 寛, 菊本弘次, 増井寛治, 益子 茂: 陽性・陰性症状評価尺度 (PANSS) 日本語版の信頼性の検討. 臨床精神医学 **22**(5): 609~614, 1993
- 8) 丸山 晋, 大島 巖, 金 吉晴, 加藤正明 訳: 精神医学的能力障害評価面接基準. 国立精神・神経センター, 精神保健研究所, 1992
- 9) 丸山 晋, 大島 巖, 加藤正明: 精神医学的能力障害の評価と低減に関する WHO の協同研究. 精神保健研究 **36**: 157~165, 1990
- 10) Wood. P. & Badley, E.M.: An epidemiological appraisal of disablement. In: Bennett, A.E., ed., Recent advances in community medicine. Edinburgh, London, and New York, Churchill Livingstone, 1978
- 11) Wing, J.K.: Impairments in schizophrenia: a rational basis for social treatment. In: Wirt, R.D. et al., Life history research in Psychopathology, Minneapolis, University of Minnesota Press, 1976
- 12) International classification of impairments, disabilities and handicaps. Geneva, World Health Organization, 1980
- 13) Ardoin M.P., Faccincani C., Galati I., Mignolli G.: Inter-rater reliability of the Disability Assessment Schedule (DAS, version II) An Italian study Soc Psychiatr Epidemiol **26**: 147~150, 1991
- 14) 氏原鉄郎: 精神分裂病に対する disability assessment schedule の臨床的意義. 慈恵医大誌 **108**: 379~390, 1993
- 15) Cohen, J.: A coefficient of agreement for nominal scales. Educational psychological measurements. **20**: 37~46, 1960
- 16) Fleiss J.L.: Measuring nominal scale agreement among many raters Psychological Bulletin **76**(5): 378~382, 1971
- 17) Bartko J.J., Carpenter W.T.: On the methods and theory of reliability. The Journal of Nervous and Mental Disease, **163**(5): 307~317, 1976
- 18) Cicchetti D.V.: Assessing Inter-Rater Reliability for Rating Scale: Resolving some Basic Issues. British Journal of Psychiatry **129**: 452~456, 1976
- 19) 北村俊則: 精神症状測定の理論と実際. 東京: 海鳴社, 1988
- 20) 友田貴子: 係数 κ をめぐる 2 つのパラドックスについて. 精神科治療学 **4**(4): 499~504, 1993
- 21) Feinstein, A.R. & Cicchetti, D.V.: High agreement but low kappa: I. the problems of two paradoxes. Journal of Clinical Epidemiology **43**: 543~549, 1990
- 22) Cicchetti, D.V. & Feinstein, A.R.: High agreement but low kappa: II. resolving the paradoxes. Journal of Clinical Epidemiology **43**: 551~558, 1990

Inter-rater Reliability of the Japanese Version of the Disability Assessment Schedule (DAS) Applied to Schizophrenics

Seiji HONGO

Department of Neuropsychiatry, Tokyo Medical College
(Director : Prof. Muneo SHIMIZU)

A disability Assessment Schedule (DAS) was developed for the “WHO collaborative study on the assessment and reduction of psychiatric disability” which aimed at establishing the criteria of psychiatric functional disturbance, psychiatric ability decrease, natural passage, and judgment of the forecast factor at the function level, and was translated into Japanese in 1991.

Double-checking of all items has not been reported, though the double check by a structured interview with 5 divisions and 61 items has been performed in many countries. In this study, double-checking was performed in 2 facilities with 32 schizophrenic patients for all items of the DAS. The κ statistic was used for statistical authorization of inter-rater reliability covering the result.

In the examination of the agreement rate between raters by the κ statistic, much higher agreement was obtained covering 46 items of all DAS item 61 item inside (the entire 75.4%). Items with low agreement rates between raters were concentrated in the 3rd division (12 of 15 items : 80.0%). This appeared to depend on “standard decentralization”. The remaining 3 items (of 15 items : 20.0%) depended on “observation decentralization”.

As a result of this study, rapid re-examination and correction in setting the anchor point and 6 step evaluation etc. of the 3rd division were indicated to be necessary. The use of DAS for tempering them was profitable in clinical applications.

〈Key words〉 Inter-rater reliability, Disability assessment schedule, Psychiatric disability, Schizophrenia, κ statistic.
