

# DSM-III 痴呆診断および柄澤式ぼけ 評価尺度の評定者間信頼度

北村俊則 丸山 晋 大塚俊男  
下仲順子 中里克治

老年精神医学 Vol.2 No.5 1985 別刷

 情報開発研究所

## DSM-III 痴呆診断および柄澤式ぼけ評価尺度 の評定者間信頼度

北村俊則\* 丸山 晋\* 大塚俊男\*  
下仲順子\*\* 中里克治\*\*

\*国立精神衛生研究所老人精神衛生部, \*\*東京都老人総合研究所心理・精神医学部

**抄録** 特別養護老人ホーム入所者52人(男性26人, 女性26人, 年齢62~97歳, 平均79.0歳)を10年以上の臨床経験をもつ精神科医3人が同席面接を行い, DSM-IIIの痴呆の各診断項目と最終判定, および柄澤式ぼけ評価尺度を独立して評点したうえで, その評定者間信頼度(一致度)を求めた. DSM-IIIの項目のなかで人格変化, 特異的器質因子に関する項目が, やや信頼度が低かったことを除けば, どの項目も満足のいく信頼度が得られ, DSM-IIIの最終判定および柄澤式ぼけ評価尺度についても満足のいく信頼度が認められた.

老年精神医学2: 774-777, 1985

**Key words**: 痴呆, 診断基準, DSM-III, 信頼度

### 序 言

痴呆に関する臨床研究を行うに際して, 痴呆患者を正確に同定し, その重症度を評価することは不可欠の手続きである. しかし, これまでのところ, 評価を行うものによって痴呆の同定に差異が生ずるかという問題に言及した研究は少ない. これは, ひとつには痴呆という臨床症状群は, 臨床家にとっては自明のものであり, また多くの評価尺度<sup>11)</sup>が存在することから, 客観的, 計量的接近を試みる時はいずれかの尺度を用いばすんでいたことによると考えられる.

しかし, 近年は, 予防的努力や薬物による治療に焦点があてられるにつれ, 軽症の痴呆が研究の対象になることが多くなった.<sup>5,9)</sup> 患者は正常者から一線を画して区別されるものではなく, 正常から

痴呆に移行するものであるから, 軽度の痴呆を研究対象とする際には, 軽度ではあっても, 病的な痴呆を加齢に伴う生理的老化から識別することが(もし可能であるならば)必要となる. また記憶力, 判断力, 計算力の低下などのほか, 人格の変化や高次皮質機能の障害など, 多種の症状が出現し, 必ずしも一次元の変化ではとらえきれないことも報告されている.<sup>3,8)</sup> したがって痴呆のどの症状を第一義的にとるかによって, 評定者による痴呆の有無の判断に差が生じることも考えられる.

精神障害の分類と各精神障害の診断方法を操作的に定義するという努力は, 1970年代から積極的に行われているが, 痴呆を操作的に規定する試み<sup>1,13)</sup>は, 1980年代に, 米国精神医学会がDSM-IIIを発表するまで, 目立ったものは認められなかった.

そこで筆者らはDSM-IIIによる痴呆診断基準の, わが国における有用性を検討する研究の一環として, その評定者間信頼度を, わが国で頻用されている柄澤式ぼけ評価尺度<sup>12)</sup>と対比しながら求め

(受付日 1985年5月13日)

Toshinori Kitamura, Susumu Maruyama, Toshio Ohtsuka, Yoshiko Shimonaka, Katsuharu Nakazato

\*〒272 市川市国府台 1-7-3

ることとした。

### 対象および方法

対象被検者は、都内の某特別養護老人ホーム入所者を無作為に抽出した。男性 26 人、女性 26 人で、年齢は 62~97 歳 (平均 79.0±8.2 歳)、入院期間は 1 か月~4 年 4 か月 (平均 2 年 11 か月 ±1 年 11 か月) であった。

結婚歴は、既婚 9 人、死別 31 人、離婚・別居 3 人、独身 5 人、不明 4 人であった。学歴は、未就学 1 人、小学校中退 3 人、小学校卒 15 人、高等小学校卒 13 人、中学校 (高等女学校、実業学校、師範学校を含む) 卒 9 人、高等学校、高等専門学校卒 3 人、大学卒 1 人、不明 7 人であった。以前の主たる職業は自営業 3 人、専門職 3 人、管理職 1 人、事務職 3 人、現業 13 人、農林漁業 1 人、サービス業 6 人、主婦 4 人、不明 18 人であった。

10 年以上の臨床経験をもつ 3 人の精神科医が 1 組となり、各被検者について同席面接を行った。1 人の精神科医が面接を施行し、面接の最後に他の 2 人の医師が必要と感じた場合は質問を追加した。その後各医師は DSM-III による痴呆の判定、および柄澤によるばけの程度の臨床的判定基準 (柄澤式ばけ評価尺度) の評点を独立して行ったうえで相互の評価を比較し、評価が異なる際は、その原因について検討を加えた。主たる面接

者と観察者の役割は各被検者によって交代した。評定の一致度は、DSM-III の項目については、Cohen の kappa ( $\kappa$ ) で、柄澤式ばけ評価尺度については、Cohen の weighted kappa でそれぞれ求めた。

DSM-III 各基準について、2 人の評定者のうち少なくとも 1 人が「あり」と評定した症例の、全症例に対する比率を base rate とし、base rate が 10% 以下、もしくは 90% 以上の場合は、 $\kappa$  の値の評価は不適切なので評価しなかった。

柄澤式ばけ評価尺度の weighted kappa を求めるにあたっては、重みづけを 2 通りにて行った。まず、2 人の評定者間の判断の相違が 1 段階ならば 1 点、2 段階ならば 2 点と重みづけを 1 点ずつ重くしていく方式 (I 式) をとり、次に、1 段階の相違は 0 点とし、2 段階に 1 点、3 段階に 2 点としていく方式 (II 式) をとった。

### 結 果

DSM-III の基準に従えば、約半数の被検者が痴呆と判定された。痴呆であるかどうかの最終的判定の評定者間信頼度は 0.62~0.84 であり平均 0.74 と高い  $\kappa$  の値が得られた (表 1)。

各項目についても C-4)、E-1)、E-2) でやや低いものの全体的には満足ゆく  $\kappa$  であるといえる。

次に、柄澤式ばけ評価尺度の評定者間信頼度に

表 1 DSM-III による痴呆の各基準の base rate と評定者間信頼度

| DSM-III の痴呆の診断基準         | 評定者(A, B, C)の組み合わせと評定者間信頼度 |            |            |
|--------------------------|----------------------------|------------|------------|
|                          | A-B                        | A-C        | B-C        |
| A. 社会的・職業的機能障害に至る知的能力の喪失 | 0.68(0.65)                 | 0.69(0.54) | 0.48(0.65) |
| B. 記憶障害                  | 0.52(0.73)                 | 0.73(0.62) | 0.42(0.70) |
| C. 1) 抽象的思考の障害           | 0.54(0.58)                 | 0.69(0.54) | 0.45(0.58) |
| 2) 判断の障害                 | 0.70(0.38)                 | 0.90(0.29) | 0.70(0.40) |
| 3) 高次皮質機能の障害             | --- (0.04)                 | --- (0.00) | --- (0.04) |
| 4) 人格変化                  | 0.50(0.19)                 | 0.30(0.17) | 0.30(0.17) |
| D. 意識混濁の欠除               | --- (1.00)                 | --- (1.00) | --- (1.00) |
| E. 1) 特異的器質因子の証拠         | 0.56(0.21)                 | 0.51(0.19) | 0.33(0.27) |
| 2) 特異的器質因子の推定            | 0.69(0.48)                 | 0.68(0.37) | 0.49(0.50) |
| 最終判定                     | 0.77(0.54)                 | 0.84(0.46) | 0.62(0.56) |

( )内は base rate.

表2 柄澤式ほけ評価尺度の評定者間信頼度

| 重みづけ方式* | 評定者(A, B, C)の組み合わせと評定者間信頼度 |      |      |
|---------|----------------------------|------|------|
|         | A-B                        | A-C  | B-C  |
| I       | 0.57                       | 0.61 | 0.50 |
| II      | 0.77                       | 1.00 | 0.91 |

\*本文参照。

関して、2種類の重みづけのどちらを使った場合も、信頼度は良好であった(表2)。さらに、各評定者の組み合わせごとの評価の内容をみると、評定者Aと評定者Cは、ほぼ類似の評価態度を示しているが、評定者Bは、評定者Aおよび評定者Cと比較して、重い評点をつける傾向があることがうかがわれた。

### 考 察

今回の調査で、DSM-IIIの痴呆評価尺度および柄澤式ほけ評価尺度は、ともに高い信頼度が示された。DSM-IIIのfield surveyはすでに米国で施行されている<sup>18)</sup>。この調査においては、老年期および初老期の痴呆のbase rateは3.1%で、同席面接による評定者間信頼度は1.00、試験再試験法では0.74という高いκ係数が示されている。米国のfield surveyにおいては、2人の評定者は同一の施設において、臨床に従事している可能性が強く、さらにDSM-IIIに習熟するため、少なくとも15人の実際の患者に適用してから、信頼度検定に進むようデザインされている。今回の調査は、痴呆のbase rateも高く、痴呆か非痴呆かの判断に困難を感じるような境界例が多かった。さらに、評価にあたった3人の医師は、同一の施設で臨床に従事した経験がなく、また1人(評定者A)を除いては、日常の臨床場面でDSM-IIIを使用したことはない。加えて、米国のfield surveyのような予備練習を行っていない。このような状況のなかで行った信頼度検定において、前述のような高いκの係数が示されたということは、DSM-IIIの痴呆概念がわが国においても困難なく使用され、診断医師間においても高い一致率が得られることを示唆している。

次に、DSM-IIIの痴呆の各項目の一致度を検討

してみると、大部分の項目で十分な信頼度が示された。しかし、人格変化と特異的器質因子の項目のκ係数が低くなっていた。人格変化は継続的な観察が不可欠であり、今回のような横断的な診

察では、評価が不安定になるものと考えられる。また、DSM-IIIの人格障害の診断の信頼度が、ほかの精神障害と比較して低いことはすでに報告されて<sup>15,19)</sup>おり、原因についても<sup>6)</sup>、使用する用語の不統一、重複診断の取り扱い、正常と障害の峻別、重症度の設定、各診断基準の操作的定義の困難性など、さまざまな検討が加えられている。そして、DSM-IIIの人格障害にあわせた自己記入式の調査票<sup>10)</sup>や、さらに構成化された面接基準<sup>16,17)</sup>も試作され、それらの妥当性の研究がなされつつある。わが国においても、痴呆に伴って出現する人格変化についても、標準化された面接基準を作成し使用することが急務であろう。さらに、家族や介護者からの情報が加われば、人格変化に関する評価の精度は向上するものと思われる。

特異的器質因子に関する判断も、神経学的考察、諸検査所見を追加すれば一致度は改善するであろう。

柄澤式ほけ評価尺度についての信頼度についての報告は、筆者らが知るかぎり今回の調査がはじめてであると思われるが、DSM-IIIと同様高い一致率が示された。

各評定者について、柄澤式ほけ評価尺度の評点態度を考察すると、評定者Bが他二者と比較して、重くつける傾向にあった。そこで、各評定者の経歴をみると、評定者Bは定期的に老人クラブにおける診察を続けているのに対し、評定者Cは、外来の診察に加えて、精神病院における老人患者を扱うことが多いことがわかった。また、評定者Aは、ごく短期間、老人病棟を受けもったことがあるが、主として成人の機能的な精神障害を扱うことが多い。このことから、評定者Bは、ごく軽度の痴呆の有無を判断することが日常の業務であることから、正常域の定義がきびしく、し

たがって、軽度の障害もぼけに含める傾向がでたのであろう。そして、評定者 A は、他二者の折衷的な評価を行おうとしたものと推測される。

このような評定者のそれまでの臨床経験<sup>2)</sup>に基づく評価の相違は、threshold variance とよばれるものである。これを修正することにより、さらに高い信頼度を得ることができる。そのためには、まず、各項目の評価についての操作的定義をさらに具体化し、評定者に標準的な訓練<sup>7)</sup>を施行することが望ましい。このような操作的標準化と訓練により、評定者の経験あるいは情報量や、その内容の相違による影響は少なくすることができる<sup>14,20)</sup>であろう。

DSM-III を用いて痴呆の診断を行うには、このようにさらに検討すべき問題が残されている。しかし、今回の研究が予備的調査であることを考慮にいれば、得られた信頼度係数は、十分満足のゆくものであり、今後とも上述したような臨床評価に関わる努力をすることにより、痴呆診断の信頼度は、さらに向上するものと考えられる。

#### 文 献

- 1) American Psychiatric Association : Diagnostic and statistical manual of mental disorders (DSM-III). 3rd ed. American Psychiatric Association, Washington D. C., 1980
- 2) Andreasen NC, McDonald-Scott P, Grove WM, Keller MB et al. : Assessment of reliability in multicenter collaborative research with a videotape approach. **Am J Psychiatry** 139 : 876-882, 1982
- 3) Ballinger BR, Reid AH, Heather BB : Cluster analysis of symptoms in elderly demented patients. **Br J Psychiatry** 140 : 257-262, 1982
- 4) Bartko JJ, Carpenter WT : On the methods and theory of reliability. **J Nerv Ment Dis** 163 : 307-317, 1976
- 5) Cooper B, Bickel H : Population screening and the early detection of dementing disorders in old age ; A review. **Psychol Med** 14 : 81-95, 1984
- 6) Frances A : The DSM-III personality disorders section ; A commentary. **Am J Psychiatry** 137 : 1050-1054, 1980
- 7) Gibbon M, McDonald-Scott P, Endicott J : Mastering the art of research interviewing ; A model training procedure for diagnostic evaluation. **Arch Gen Psychiatry** 38 : 1259-1262, 1981
- 8) 原田憲一 : 老人における痴呆の臨床類型. **精神経誌** 83 : 117-128, 1981
- 9) Henderson AS, Juppert FA : The problems of mild dementia. **Psychol Med** 14 : 5-11, 1984
- 10) Hyler SE, Rieder RO, Spitzer RL, Williams JBW : Personality questionnaire. New York State Psychiatric Institute, New York, 1983
- 11) 伊藤 斉, 武井妙子, 立山万里, 神定 守 : 痴呆の評価尺度 ; 治療効果の評価への応用について. **臨床評価** 11 : 527-554, 1983
- 12) 柄澤昭秀 : 老人のぼけの臨床. 医学書院, 東京, 1981
- 13) Lipowski ZJ : Organic mental disorders ; An American perspective. **Br J Psychiatry** 144 : 542-546, 1984
- 14) McDonald-Scott P, Endicott J : Informed versus blind ; The reliability of cross-sectional ratings of psychopathology. **Psychiatry Res** 12 : 207-217, 1984
- 15) Mellsop G, Varghese F, Joshua S et al. : The reliability of axis II of DSM-III. **Am J Psychiatry** 139 : 1360-1361, 1982
- 16) Pfohl B, Stangl D, Zimmerman M : Structured interview for the DSM-III personality disorders (SIDP). 3rd ed. Department of psychiatry, University of Iowa, Iowa city, 1983
- 17) Pfohl B, Stangl D, Zimmerman M : The implication of DSM-III personality disorders for patients with major depression. **J Affective Disord** (in press)
- 18) Spitzer RL, Forman JBW, Nee J : DSM-III field trials ; Initial interrater diagnostic reliability. **Am J Psychiatry** 136 : 815-817, 1979
- 19) Spitzer RL, Forman JBW : DSM-III field trials ; II. Initial experience with the multiaxial system. **Am J Psychiatry** 136 : 818-820, 1979
- 20) Webb LJ, Gold RS, Johnstone EE, Diclemente CC. : Accuracy of DSM-III diagnoses following a training program. **Am J Psychiatry** 138 : 376-378, 1981