

Table 1 The expression of bHLH transcription factors after chronic treatment with antidepressant in rat frontal cortex as determined by real time RT-PCR. Total RNA was extracted from frontal cortex of rats treated with vehicle, imipramine, sertraline or fluoxetine for 21 days. Total RNA was used for real time quantitative PCR.

	imipramine (10 mg/kg)	sertraline (10 mg/kg)	fluoxetine (5 mg/kg)
Mash1	↑	→	→
Math1	N.A.	N.A.	N.A.
Neurogenin	N.A.	N.A.	N.A.
Math2	↑	↑	↑
NeuroD1	→	→	→
NeuroD2	↓	↓	↓

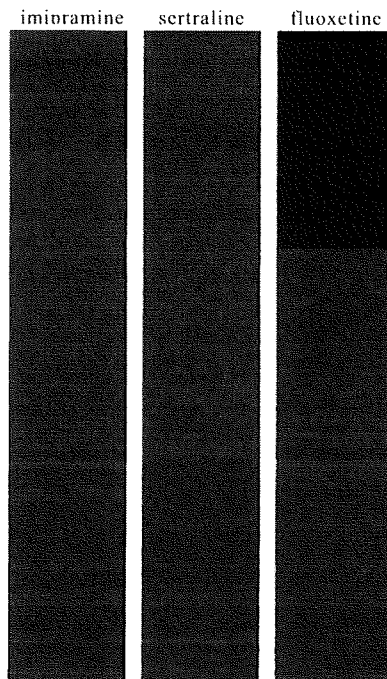


Fig. 1 Analysis of gene expression patterns subjected with antidepressant treatment in rat frontal cortex

Total RNA was extracted from frontal cortex of rats treated with vehicle, imipramine, sertraline or fluoxetine for 21 days. Total RNA was used for GeneChip analysis. In each sample, up-regulated genes are colored in red and down-regulated genes are in green when compared to the control.

トコールに基づき、施設内審査委員会の規定に則り倫理性と科学性に十分配慮して実施した。

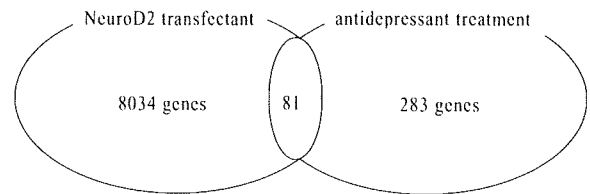


Fig. 2 Classification of genes on the GeneChip array

The expression of 8034 genes were induced or reduced in the NeuroD2 transfectant in rat cortical neuronal cultured cells, 283 genes were changed in rat treated with antidepressant analyzed by GeneChip. After comparison of these genes, 81 genes were found in common.

結 果

抗うつ薬投与による bHLH 型転写因子の発現の定量：

抗うつ薬長期投与により発現変化する遺伝子の網羅的解析により、発現変化が認められた ADRG#701 は、bHLH 構造を有する転写因子であることが明らかとなった。bHLH 構造を有する転写因子は他にも複数存在し、ファミリーを形成している。そこで、これらを含めた bHLH 型転写因子 (Mash1, Math1, Neurogenin, NeuroD1, NeuroD2, Math2) について特異的プライマーを設計し、real time RT-PCR 法により抗うつ薬投与における発現量を定量した。その結果、Math2 は imipramine, sertraline, fluoxetine の 3 種の

抗うつ薬により共通して発現増加すること、NeuroD2は発現減少することが明らかとなった (Table 1)。一方、Mash1, NeuroD1は3種の抗うつ薬による共通した発現変化は認められなかった (Table 1)。また、Math1, Neurogeninはadult rat 前頭葉皮質では発現が非常に少なく、定量は出来なかった (Table 1)。

GeneChip 解析による NeuroD2 の下流ターゲットの探索：

次に、抗うつ薬長期投与により発現変化することが明らかとなった NeuroD2 が転写を制御する下流ターゲット遺伝子の探索を行うため、大脳皮質初代培養神経細胞に NeuroD2 をトランスフェクションし、過剰発現する細胞の調製を行った。しかし、同細胞へのトランスフェクション効率が非常に低かったため (約5%未満)、トランスフェクションされた細胞のみを分離・濃縮することを試みた。コントロール、NeuroD2 を組み込んだ発現ベクター及びトランスフェクションされた細胞を分離するための表面抗原 LNGFR を co-transfection し、24 時間後、磁気ビーズを付加した抗 LNGFR 抗体を用いてトランスフェクションされた細胞のみを分離した。その結果、トランスフェクションされた細胞は分離操作後に得られた全細胞の約 80% に向上した。そこで、この細胞を用いて GeneChip による発現プロファイル解析を行った。発現が確認されたスポットは 18,528 遺伝子であり、コントロールに比べて NeuroD2 発現細胞で up-regulation していた遺伝子が 5,178 遺伝子、down-regulation していた遺伝子が 2,856 遺伝子であった (Fig. 1)。また、imipramine, sertraline, fluoxetine 長期投与ラット前頭葉皮質を用いて同 GeneChip を解析した結果、3種の抗うつ薬で共通して up-regulation していた遺伝子が 83 遺伝子、down-regulation していた遺伝子が 200 遺伝子であった (Fig. 1, 2)。さらに、NeuroD2 発現細胞及び3種の抗うつ薬により共通して発現変化する遺伝子は 81 遺伝子であった (Fig. 2)。これらの遺伝子について機能別分類を行ったところ、神経細胞の成熟や軸索形成、細胞間コミュニケーション、細胞内応答等に関与する遺伝子が含まれていた。

考 察

これまで「抗うつ薬投与による転写因子 CREB のリン酸化亢進とこれに続く BDNF 等のターゲット遺伝子の転写調節、BDNF による機能変化 (神経保護、神経新生)」等の一連の流れがうつ病の治癒につながるという報告が多数なされ、うつ病治癒メカニズム仮説として一定の評価を受けている。しかし BDNF の増加がすべての抗うつ薬に共通して認められる現象ではないことなど、この仮説でうつ病治癒メカニズムのすべてを説明するには至っていない¹⁾。一方、これまでに我々は、網羅的スクリーニングによりうつ病の治癒機転に重要な遺伝子として antidepressant related gene (ADRG) を探索してきたが、その中に bHLH 構造を有する転写因子を同定した。

bHLH 構造を有する転写因子は、発生の時期に対応し中枢神経系の部位特異的に発現する。初期に発現する Hes1, Hes5 は、その欠損したマウスにおいて神経幹細胞の維持が不完全で正常より早く分化が起こること、また、これらを持続発現させるとニューロンへの分化が抑制され未分化状態が維持されることから、Hes1, Hes5 は、神経幹細胞の維持に重要な役割を担っていることが報告された⁴⁾。また、Hes1, Hes5 は Mash1, Math1, Neurogenin の発現を負に制御しており、Hes1, Hes5 の発現減少に伴い、Mash1, Math1, Neurogenin の発現が上昇する。Mash1, Math1, Neurogenin は、ニューロンへの分化の決定に必須であることが明らかとなった^{5,6)}。これらの転写因子により発現が正に制御されている Math2, NeuroD1, NeuroD2 は、神経突起を正しい方向に伸展させる機能を有していることが報告されている⁴⁾。

本研究では、抗うつ薬長期投与による Mash1, Math1, Neurogenin, Math2, NeuroD1, NeuroD2 のラット大脳皮質における発現変化を定量した。Real time RT-PCR の結果、Math2 は imipramine, sertraline, fluoxetine の3種の抗うつ薬により共通して発現増加すること、NeuroD2 は発現減少すること、NeuroD1 は共通した発現変化が認められないことが明らかとなった。Math2,

NeuroD1, NeuroD2 は相同性が高く、脳の発達段階においては背側感覚上皮に高発現し、成体期まで継続して大脳皮質、海馬、小脳に発現しているというように発現パターンが類似していることが報告されている⁷⁾。しかしながら、NeuroD1 欠損マウスは膵臓の機能的欠陥により糖尿病になり生後まもなく死亡すること⁸⁾、NeuroD2 欠損マウスでは、体型が小さく未成熟のまま死亡すること、また、運動失調、痙攣閾値の減少などが認められることなど⁹⁾、表現系が異なることが報告されている。Math2, NeuroD1, NeuroD2 の抗うつ薬に対する発現変化の差異については今後さらなる検討が必要である。

次に、NeuroD2 がマスター遺伝子として転写を制御するターゲット遺伝子群を探索するため、NeuroD2 を過剰発現した大脳皮質初代培養神経細胞を調製し、GeneChip による遺伝子発現プロファイル解析を行った。さらに、imipramine, sertraline, fluoxetine を長期投与したラット前頭葉皮質で共通して発現変化する遺伝子群と比較検討した。その結果、NeuroD2 発現細胞及び3種の抗うつ薬により共通して発現変化する遺伝子81遺伝子を同定し機能別分類を行ったところ、神経細胞の成熟や軸索形成、細胞間コミュニケーション、細胞内応答等に関与する遺伝子が含まれていた。

NeuroD2 を P19 (embryonal carcinoma cells) に過剰発現させると神経特異的タンパクの発現が増加し、さらに神経様の形態を呈することが報告されている¹⁰⁾。NeuroD2 欠損マウス (heterozygote) は、恐怖条件付け試験において、emotional learning が減少することが明らかとなった¹¹⁾。これらのマウスでは、扁桃体で synaptic remodeling に関わる AMPA 及び GABA_A 受容体が減少していることが報告されている¹¹⁾。本研究で同定した NeuroD2 の下流ターゲット遺伝子の機能別クラスタリングの結果及びこれらの報告を考え合わせると、転写因子である NeuroD2 により制御される下流ターゲット遺伝子がニューロン及び受容体に影響を及ぼし、synaptic remodeling 等の機能変化を介し感情調節に関わる可能性が考えられた。

今後、NeuroD2 がマスター遺伝子として転写

を制御する個々のターゲット遺伝子について検討をしていくことにより、「CREB-BDNF-神経可塑的变化」にかわる新たなうつ病治療仮説が提示できるものとする。

文 献

- 1) Malberg JE, Blendy JA: Antidepressant action: to the nucleus and beyond. *Trends Pharmacol Sci* 26 : 631 - 638, 2005
- 2) Yamada M, Yamada M, Kiuchi Y, Nara K, Kanda Y, Morinobu S, Momose K, Oguchi K, Kamijima K, Higuchi T: Identification of a novel splice variant of heat shock cognate protein 70 after chronic antidepressant treatment in rat frontal cortex. *Biochem Biophys Res Commun* 261 : 541 - 545, 1999
- 3) Yamada M, Yamada M, Yamazaki S, Takahashi K, Nishioka G, Kudo K, Ozawa H, Yamada S, Kiuchi Y, Kamijima K, Higuchi T, Momose K: Identification of a novel gene with RING-H2 finger motif induced after chronic antidepressant treatment in rat brain. *Biochem Biophys Res Commun*, 278 : 150 - 157, 2000
- 4) Kageyama R, Ohtsuka T, Hatakeyama J, Ohsawa R: Roles of bHLH genes in neural stem cell differentiation. *Exp Cell Res* 306 : 343 - 348, 2005
- 5) Ma Q, Kintner C, Anderson DJ: Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* 87 : 43 - 52, 1996
- 6) Sommer L, Ma Q, Anderson DJ: Neurogenins, a novel family of atonal-related bHLH transcription factors, are putative mammalian neuronal determination genes that reveal progenitor cell heterogeneity in the developing CNS and PNS. *Mol Cell Neurosci* 8 : 221 - 241, 1996
- 7) Schwab MH, Druffel-Augustin S, Gass P, Jung M, Klugmann M, Bartholomae A, Rossner MJ, Nave KA: Neuronal basic helix-loop-helix proteins (NEX, neuroD, NDRF): spatiotemporal expression and targeted disruption of the NEX gene in transgenic mice. *J Neurosci* 18 : 1408 - 1418, 1998
- 8) Naya FJ, Stellrecht CM, Tsai MJ: Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. *Genes Dev* 9 : 1009 - 1019, 1995
- 9) Olson JM, Asakura A, Snider L, Hawkes R, Strand A, Stoeck J, Hallahan A, Pritchard J, Tapscott SJ: NeuroD2 is necessary for development and survival of central nervous system neurons. *Dev Biol*. 234 : 174 - 187, 2001

- 10) Farah MH, Olson JM, Sucic HB, Hume RI, Tapscott SJ, Turner DL: Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. *Development* 127 : 693 - 702, 2000
- 11) Lin CH, Hansen S, Wang Z, Storm DR, Tapscott SJ, Olson JM: The dosage of the neuroD2 transcription factor regulates amygdala development and emotional learning. *Proc Natl Acad Sci USA* 102 : 14877 - 14882, 2005

ABSTRACT

Identification of candidate genes regulated by transcription factor related with therapeutic mechanism for depression

Mitsuhiko Yamada*, Kou Takahashi*, Yoshiko Shida*, Yoshiaki Maruyama*, Misa Yamada*, Teruhiko Higuchi**

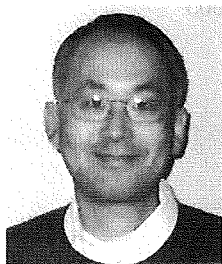
* Division of Psychogeriatrics, National Institute of Mental Health, National Center of Neurology and Psychiatry; 4-1-1, Ogawahigashimachi, Kodaira, Tokyo 187-8553, Japan. ** Musashi Hospital, National Center of Neurology and Psychiatry; 4-1-1, Ogawahigashimachi, Kodaira, Tokyo 187-8553, Japan.

While performing differential cloning study to identify some molecular machinery responsible for the antidepressant effect, we found a bHLH transcription factor as a candidate. Using real time quantitative RT-PCR, we demonstrated that chronic treatment with three antidepressants, imipramine, sertraline and fluoxetine significantly reduced NeuroD2 mRNA expression in rat frontal cortex. To elucidate the NeuroD2 transcription network mediated by antidepressant, we used GeneChip analysis in rat cortical neuronal cells transfected with NeuroD2. The expression of 8034 genes were induced or reduced in the NeuroD2 transfectant in cortical neuronal cells. We also analyzed the sample of rats treated with antidepressant. The expression of 283 genes were commonly changed in imipramine, sertraline and fluoxetine treated rat frontal cortex. After comparison of these genes changed by NeuroD2 transfection and antidepressant treatment, 81 genes were found in common. These genes were classified to 1) neuronal maturation related genes, 2) axon guidance related genes, 3) cell-cell communication and cell response related genes. These functional changes may be related to the therapeutic mechanism of depression.

自殺対策のための戦略研究

— J-MISP

Japanese Multimodal Intervention Trials for Suicide Prevention (J-MISP)



山田光彦(写真) 高橋清久

Mitsuhiro YAMADA¹ and Kiyohisa TAKAHASHI²

国立精神・神経センター精神保健研究所老人精神保健部¹, (財)精神・神経科学振興財団²

◎厚生労働科学研究費補助金こころの健康科学研究事業により、“自殺対策のための戦略研究”が2005年度より開始された。本戦略研究では医療モデルを超えた複合的自殺対策の立案をめざしている。実施主体は(財)精神・神経科学振興財団であり、国立精神・神経センターが平常的かつ専門的な研究支援を行っている。具体的には2つの比較研究“複合的自殺対策プログラムの自殺企図予防効果に関する地域介入研究(NOCOMIT-J)”および“自殺企図の再発防止に対する複合的ケースマネジメントの効果：多施設共同による無作為化比較研究(ACTION-J)”を準備し、2006年度よりすでに開始している。



Key word

自殺, 戦略研究, 無作為化比較試験, 地域介入, ケースマネジメント

わが国では1998年に年間自殺者が前年度比130%以上という、他国に類のない激増をみた。現在でも実に交通事故による死者数の約4~5倍の人が毎年自殺によって命を落としている。このようななかで、自殺防止対策有識者懇談会は“自殺予防に向けての提言”を2002年に報告しており¹⁾、社会全体として自殺に取り組むことが提言された。このように、わが国の社会において自殺問題はきわめて深刻な問題であり、その対策は急務である。実際、わが国では年間自殺者数が3万人を超えている。さらに、自殺未遂はその10倍以上ともいわれており、自殺や自殺未遂によって家族や友人など周囲の人びとが受ける心理的影響を考慮すると、毎年、百数十万人の人びとが自殺問題に苦しんでいることになる。

2005年度より、自殺対策のための実証的根拠を得ることを目的として厚生労働科学研究費補助金こころの健康科学研究事業により“自殺対策のための戦略研究(Japanese Multimodal Intervention

Trials for Suicide Prevention: J-MISP)”が開始された(表1)²⁾。本戦略研究では“自殺対策”を最終目標とし、医療モデルを超えた複合的な対策の立案をめざしている。本稿では、この研究プログラムが開始された背景とその全体像について紹介することを試みる。

戦略研究とは

欧米では、保健医療福祉システムや疾病予防/治療法の効果を実証的に評価することができる“無作為化比較試験”、あるいは多数の患者の医療情報を長期間にわたって集積・解析し疾患の実態を明らかにする“観察研究”が、これまでに数多く実施されてきた。その結果、“数多くの客観的な証拠”に基づいて各種疾患の予防・診断・治療の標準化が試みられ、“理論や経験に依存した医療”から“実証的根拠に基づく医療”への転換がはかられてきた。しかし、わが国ではこうした臨床研究への取組みが大きく立ち遅れており、全国的かつ継続

表 1 自殺対策のための戦略研究

統括責任者(厚生労働科学研究費補助金主任研究者):(財)精神・神経科学振興財団理事長 高橋清久
 運営管理:国立精神・神経センター精神保健研究所部長 山田光彦
 事務ロジスティック担当:(財)精神・神経科学振興財団戦略研究担当事業部

(1) 複合的自殺対策プログラムの自殺企図予防効果に関する地域介入研究(A community intervention trial of multimodal suicide prevention program in Japan : A Novel multimodal Community Intervention program to prevent suicide and suicide attempt in Japan : NOCOMIT-J)

研究リーダー:大野 裕 慶應義塾大学教授(精神医学)
 サブリーダー:酒井明夫 岩手医科大学教授(精神医学)
 事務局長:大塚耕太郎 岩手医科大学講師(精神医学)

(2) 自殺企図の再発防止に対する複合的ケースマネジメントの効果——多施設共同による無作為化比較研究(A randomized, controlled, multicenter trial of post-suicide attempt case management for the prevention of further attempts in Japan : ACTION-J)

研究リーダー:平安良雄 横浜市立大学教授(精神医学)
 研究顧問:有賀 徹 昭和大学教授(救急医学)
 事務局長:河西千秋 横浜市立大学准教授(精神医学)

的な調査もきわめて少ない。こうした現状はすでに日本学術会議の報告(2004年5月20日)“わが国における臨床疫学研究推進のための基盤整備について”にも示されているところである³⁾。

戦略研究とは、わが国を支える多くの国民の健康を維持・増進させるために、優先順位の高い慢性疾患・健康障害を標的として、その予防・治療介入および診療の質改善介入など、国民の健康を守る政策に関連する“実証的根拠(エビデンス)”を生み出すために実施される大型の臨床研究である。2005年度より創設された戦略研究は、“厚生労働省が、あらかじめ国民のニーズに基づいて策定された行政の方針に従って具体的な政策目標を定めようとして、成果(アウトカム)指標と研究計画の骨子を定める”という点で、成果指標・研究計画をすべて研究者に一任してきたこれまでの厚生労働科学研究の一般公募研究あるいは班研究とは一線を画すものである。戦略研究の成果指標および研究計画の骨子は、その研究成果を“政策”として全国に均てん化することを前提として作成されなければならない。つまり戦略研究はその成果が“現実的な問題解決”のために利用されることを前提として実施される研究なのである。

一方、近年の自殺死亡者数増加の背景には健康問題(精神疾患・身体疾患)、経済・生活問題、家庭問題のほか、人生観・価値観や地域・職場・学校教育のあり方の変化など、さまざまな社会的要

因が複雑に関係しており、予防対策の実施にあたっては多角的な検討と包括的な対策が必要になる。しかし、効果的な複合的自殺予防対策のあり方に注目した研究および施策ははなはだ不十分な現状である。そのため、全国各地の先駆的な取り組みの経験を踏まえ、大規模多施設共同研究で効果的な支援方法に関するエビデンスを構築して今後の政策立案に役立てることが必要である。具体的には“地域特性に応じた複合的自殺予防プログラムの開発”“自殺企図者の再発防止策の開発”が必要であり、自殺者数の減少に向けた取り組みが重要かつ緊急の課題として必要と考えられた。そこで、このような背景のもとに2005年度より開始される2つの戦略研究課題として、“糖尿病予防のための戦略研究”とともに、“自殺対策のための戦略研究”が選定された。

“自殺対策のための戦略研究”の実施体制

本戦略研究の実施主体は、主任研究者(高橋清久)が理事長を務める(財)精神・神経科学振興財団であり、国立精神・神経センターが高度先進的な研究機関として平常的かつ専門的な研究支援を行っている。

具体的には2005年度から2つの試験研究、“複合的自殺対策プログラムの自殺企図予防効果に関する地域介入研究(NOCOMIT-J)”および“自殺企図の再発防止に対する複合的ケースマネジメント

トの効果：多施設共同による無作為化比較研究(ACTION-J)”の準備を進めて、2006年度よりすでに開始している(表1)。以下に、2つの試験研究プロジェクトについて順に紹介する。

“複合的自殺対策プログラムの自殺企図予防効果に関する地域介入研究(NOCOMIT-J)”のおもな研究目的は、自殺死亡率が長年にわたって高率な地域において、一次から三次(ポストベンション)までのさまざまな自殺予防対策を組み合わせた新しい複合的自殺予防対策プログラム(以下、自殺対策プログラムと略)を介入地区で実施し、通常自殺予防対策を行う対照地区と比較して、自殺企図(自殺死亡および自殺未遂)の発生に効果があるかどうかを検討することである。さらに、近年急激に自殺者数が増加した大都市圏において有効な自殺予防対策を確立するため、人口が密集している都市部地域において新しい自殺対策プログラムを実施し、自殺企図(自殺死亡および自殺未遂)の発生に効果があるのかも並行して副次的に検討することである。二次予防対策においてスクリーニングされたうつ病/自殺ハイリスク者には適切な精神医学的対応が不可欠である。また、自殺予防対策という三次予防とは、自死遺族が近親者の自殺を自らの責任であるかのようにとらえたり隣人や地域との交流を閉ざしてしまわないように配慮し、グリーフワーク(喪の作業)を支えるケアを提供することを指す。このプロジェクトは、さまざまな行政機関・医療機関・教育機関やNPOなどの組織、そして地域住民のみなさんの理解と協力なしには研究を進めることができない大規模な地域介入研究である。

“自殺企図の再発防止に対する複合的ケースマネジメントの効果：多施設共同による無作為化比較研究(ACTION-J)”のおもな研究目的は、救急医療施設に搬送された自殺未遂者に対して精神科的評価および心理教育を行い、その後に試験介入としてケースマネジメントを行い、試験介入が通常介入と比較して自殺企図再発の防止に効果があるかどうかを検証することである。主要評価項目は自殺企図(自殺既遂および未遂)の再発としている。このプロジェクトは救急医療施設に搬送された自殺未遂者に対する現実的で効果的な支援

法の開発をめざした目標対象者数が1,000名を超える大規模多施設共同研究である。

自殺対策基本法の成立を受けて

2006年6月15日に“自殺対策基本法”が成立した¹⁾。この法律の目的は、自殺対策を総合的に推進して自殺の防止をはかり、あわせて自殺者の親族などに対する支援の充実をはかり、もって国民が健康で生きがいをもって暮らすことのできる社会の実現に寄与することとされている。著者らは、自殺予防に携わるすべての個人や組織がそれぞれの立場で考え、実際に行動し、つながりあい、そして学ぶことがすべての自殺対策の出発点だと考えている。“自殺対策基本法”により“自殺対策のための戦略研究”実施のための基盤がより明確なものとなった。

ここで、国の事業としての自殺対策についてすこし考えてみたい。たとえば、集中豪雨の後に河川が氾濫しいまにも堤防が決壊しようとしているとき、皆で協力して土嚢を積むことによって被害を未然に防ぐことができるかもしれない。しかし、本当の原因である“上流の森林開発”についての理解や、その川の“生態系の変化”などを十分に考慮しないかぎり、“100年もつ頑丈な堤防”を適切に設計することはできない。自殺対策についても同様である。何においてもいますぐやれる施策を迅速に実施することはとても大事だ。しかし、医療や保健福祉をよりよくしていくためには、“その介入方法に効果があるのか”“その対策が社会にどのような予想外の影響を与えるのか”などコツコツと検証し、科学的な証拠を積み上げていくしかない。

たしかに、すべての自殺を防止することは不可能な試みであるかもしれない。しかし、世の中に“あつてもよい自殺”などないのである。著者らが“自殺対策のための戦略研究”を通じて真に伝えるべきは、“自殺は避けることができる”というメッセージである。この理念があればこそ、自殺対策が“持続性をもった当たり前の行政施策”として根をおろしていくものであると考えている。

おわりに

“本人が自殺しようとしているのだから止めることはできない”という意見を聞く機会も少なくない。しかし、これは大きな誤解である。多くの場合、自殺は避けることのできる死なのである。また、自殺の背景は複雑であり、生物・心理・社会的要因を十分に検討した複合的自殺対策プログラムの重要性が指摘されている。著者らは、“自殺対策基本法”を強固な足場として“自殺対策のための戦略研究”を実施していくことにより、わが国で自殺対策を進めていくうえでの実証的根拠を

提供していくことができればと期待している。

URL

- 1) 自殺対策基本法，自殺防止対策有識者懇談会：自殺予防に向けての提言。 <http://www.ncnp.go.jp/iki-ru-hp/> (他の資料は自殺予防対策支援ページ“いきる”より入手可能)
- 2) 自殺対策のための戦略研究ホームページ： <http://www.jfnm.or.jp/itaku/J-MISP/index.html>
- 3) 日本学会議：わが国における臨床疫学研究推進のための基盤整備について。 <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-19-t1013.pdf>

* * *

分子精神医学

別刷

発行：株式会社 先端医学社
〒103-0007 東京都中央区日本橋浜町 2-17-8 KDX浜町ビル

人を対象とした医学研究を行うときに まず考えるべきこと

米本直裕*§ 中井亜弓* 山田光彦*

* 国立精神・神経センター精神保健研究所老人精神保健部
§ 京都大学大学院医学研究科社会健康医学系専攻医療統計学分野

いま、あなたは精神医学領域の研究者で、人を対象とする医学研究を開始しようと考えている。あなたはまず何をすべきだろうか？人を対象とする医学研究では、研究対象者の権利を保護し、公正な研究を行うためのいくつかの「ルール」を守らなければならない。そのルールの基本が「ヘルシンキ宣言」¹⁾である。

人を対象とする研究にはいくつかの種類がある。図1に、人を対象とする医学系研究とそれぞれに対応する指針などを示した。医薬品・医療機器の製造販売承認、保険適用を目的とした企業主導もしくは医師主導治験は薬事法の規制下であり、Good Clinical Practice (GCP)²⁾を遵守し、ICH ガイドライン³⁾に従って進める必要がある。それ以外のいわゆる研究者主導型臨床試験、疫学研究などでは、「臨床研究に関する倫理指針」⁴⁾、「疫学研究に関する倫理指針」⁴⁾を遵守して研究を進める必要がある。また遺伝子情報を取り扱う場合は、「ヒトゲノム・遺伝子解析研究に関する倫理指針」に従う必要がある⁴⁾。

疫学研究に関する倫理指針は、その対象を「人の疾病の成因及び病態の解明並びに予防及び治療の方法の確立を目的とした疫学研究」とし、これに携わるすべての関係者に遵守を求めている。臨床研究に関する倫理指針は、臨床研究の定義を「医療における疾病の予防方法、

診断方法及び治療方法の改善、疾病原因及び病態の理解並びに患者の生活の質の向上を目的として実施される医学系研究であって、人を対象とするもの（個人を特定できる人由来の材料及びデータに関する研究を含む）」とする。「医学系研究」には「歯学、薬学、看護学、リハビリテーション学、予防医学、健康科学に関する研究が含まれる」とあり、こころの健康を扱い介入を行う心理学的研究や精神療法の評価方法の妥当性に関する研究もこの範囲に含まれる。臨床研究に該当しないものとしては、① 診断及び治療のみを目的とした医療行為、② 他の法令及び指針の適用範囲に含まれる研究、としている。一方、この指針の適用範囲でよく問題となるのが、どこまでが「医療行為」でどこからが「研究」なのか不明確な点である。細則にはあくまで医療行為の一部である例として、病院内での症例検討会のような「特定の患者の治療を前提とせず、ある疾病の治療方法等を検討するため」のものをあげているが、境界を明確に線引きすることは難しいのが実情である。そのため、院外の不特定多数に該当資料が公開される可能性がある場合には、つねに「研究」であると考えて慎重に準備を進める必要があるだろう。

人を対象とする研究を行ううえでの倫理的な配慮に関するポイントは、インフォームド・コンセント、倫理審

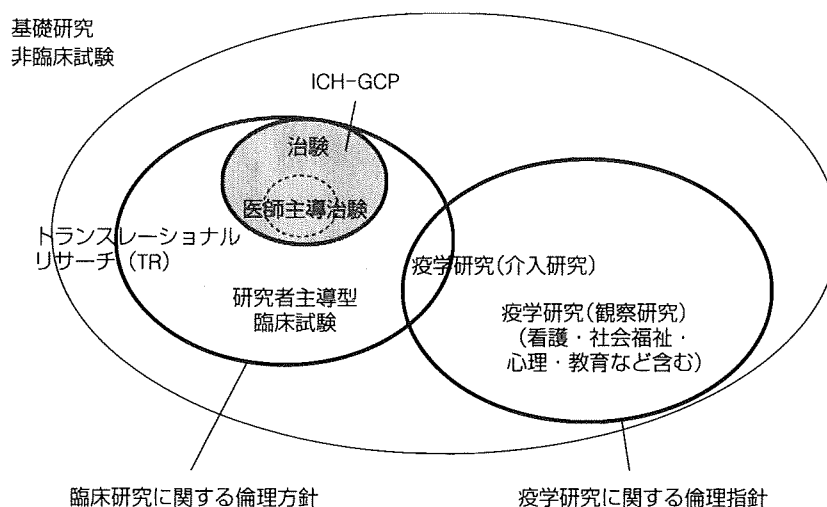


図 1. 人を対象とする医学系研究

査委員会による審査，個人情報保護の3つがあげられる。

インフォームド・コンセントは研究参加者本人から事前に書面で受けることが原則である。インフォームド・コンセントにおいて説明をしなければならない項目は，各指針の「インフォームド・コンセントの受領に関する細則」に示されている。説明文書にはここに示された項目がすべて記載される必要がある。また学術的な専門用語などに配慮して，誰にでもわかりやすく説明が行われなければならない。本人からの同意が難しい場合に，代諾者からインフォームド・コンセントを受ける場合もある。ただし，公衆衛生の向上のために特に必要で，研究参加者の協力により研究を実施することが必要不可欠であり，倫理審査委員会の承認を得て，研究機関の長の許可を受けたときに限り認められている。また，未成年，その他の行為能力がないとみられる研究参加者が，研究への参加の決定を理解できる場合には，代諾者のみならず研究参加者本人にも理解を得なければならない。未成年者の場合には，わかりやすい言葉で十分な説明を行い，理解が得られるよう努めなければならない。16歳以上の場合には，代諾者ととも本人からのインフォームド・コンセントを受ける必要がある。

地域介入研究のような「集団単位で行う介入研究」の場合では，インフォームド・コンセントを受けることは必ずしも必要ではなく，そのかわりに研究者は研究に関

する情報を公開し，研究対象となる者が研究対象となることを拒否できるようにしなければならない，と疫学研究に関する倫理指針は定めている。

人に関する研究を行うにあたっては，研究の科学的・倫理的妥当性が記載された研究計画書を作成しなければならない。研究計画書には，研究の背景，目的，デザイン，期間，対象者の選定方針（選択基準と除外基準），データ収集方法と個人情報保護，研究参加予定者数とその根拠，統計解析，インフォームド・コンセントの手続きと各文書（説明文書・同意文書），研究参加の利益と不利益，起こりうる危険ならびに不快な状態，資金源や起こりうる利害の衝突の開示，参加するすべての研究機関の名称，研究者の氏名などを記載する必要がある。倫理審査が必要な場合，研究者は研究を実施するすべての施設の倫理審査委員会に研究計画書を提出し，審査，承認を受ける必要がある。

研究を通じて集められた個人情報は，十分な安全管理措置をとり，情報の漏洩，混交，盗難，紛失が起こらないように管理する義務がある。また，研究参加者から情報開示を求められた場合には，すみやかに対応できるようにしなければならない。

他の機関への情報の提供に関しては，その資料が匿名化されることが必要である。匿名化とは，「個人情報から個人を識別できる情報の全部又は一部を取り除き，代わりにその人と関わりのない符号，または番号を付す」

ことである。「資料に付随する情報のうち、ある情報だけでは特定の人を識別できない情報であっても、各種の名簿等での他の入手できる情報との組み合わせることにより、その人を識別できる場合には、組み合わせに必要な情報の全部又は一部を取り除いて、その人が識別できないよう」する必要がある。ここでいう匿名化とは、「(個人情報との) 連結不可能匿名化または連結可能匿名化であって(個人情報との) 対応表を有していない場合」である。もちろん、研究結果を公表する場合には、個々の研究参加者を特定できないようにしなければならない。

人を対象とする研究を行う際には、まず該当する倫理指針のすべてに目を通し、実施する研究計画に問題がないか、実施体制が整っているかをチェックする必要がある。各指針には別添としてQ & Aがあり、個々の研究上の疑問に答えており、参考となる。また指針内容は改訂されるので、ホームページでつねに最新の情報を把握する必要がある(アドレスは文献を参照)。現在、疫学研究に関する倫理指針は、厚生科学審議会科学技術部会「疫学研究指針の見直しに関する専門委員会」において、2007年の改訂を目指して議論が続けられている。

倫理指針はあくまで研究のガイドであって、最低限守るべきことであり、決してそれさえ守れば十分というわ

けではない。各指針には精神領域の研究で特に生じうる倫理的問題にはほとんど触れられていない。また、人を対象とする医学研究を実施する大前提として、医療の進歩を目的とした十分な合理性のある仮説が必要であることはいまでもない。研究を行う意味が本当にあるのかがまず初めに問われるべきである。なお、質の高い疫学研究や臨床研究を計画し実行するためには、研究の計画の段階から疫学・生物統計学や医療倫理学などの専門家へ相談することをお勧めしたい。



文 献

- 1) ヘルシンキ宣言(日本医師会訳): http://www.med.or.jp/wma/helsinki02_j.html
- 2) GCP関係(日本医師会 治験促進センター): <http://www.jmacct.med.or.jp/notification/index.html>
- 3) ICHガイドライン(医薬品医療機器総合機構): http://www.pmda.go.jp/ich/ich_index.html
- 4) 厚生労働省: 医学研究に関する指針一覧
臨床研究に関する倫理指針
疫学研究に関する倫理指針
ヒトゲノム・遺伝子解析研究に関する倫理指針
<http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/index.html>

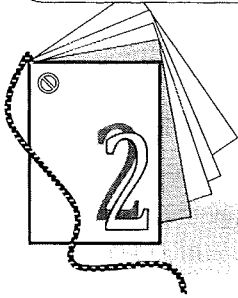
基礎講座

老年精神医学研究の進め方と発表の仕方②

研究デザインの妥当性

米本直裕・山田光彦

基礎講座：老年精神医学研究の進め方と発表の仕方



研究デザインの妥当性

国立精神・神経センター精神保健研究所老人精神保健部 米本直裕, 山田光彦

はじめに

人間を対象とする研究では、細胞や動物での実験とは異なり、倫理的かつ物理的な限界から、繰り返し実験を行うことがむずかしい。なかには、観察のみが可能で試験的な介入を行うことさえむずかしい場合もある。また、研究対象者への介入には限界があり、対象者へのリスクを考え、仮説から検証に向かって段階的に研究を進める必要がある。さらに、ヒトでは、測定指標の個体間、個体内のばらつきも大きく、環境要因などの影響も無視できない。老年精神医学研究もその例外ではない。したがって、より真実に近い研究結果を得るためには、当初から、誤差 (error) の存在を認識し、この誤差をなるべく少なくするためのさまざまな工夫を研究計画に盛り込んでおく必要がある。

I. 偶然誤差と系統誤差

誤差には大きく分けて2つの種類がある。1つは偶然誤差 (random error)、もう1つは系統誤差 (systematic error) である。偶然誤差とは標的となる母集団からサンプリングする際に起こる誤差であり、通常、その研究における対象者数と対象者のばらつきによって影響される。系統誤差は、サンプリングにおいて、何らかのある要因が引き起こす偏りであり、バイアス (bias) と呼ばれるものである。

偶然誤差と系統誤差の違いを簡単に説明してみよう。研究の対象者数を無限大に増やしたとして、ゼロになる誤差が偶然誤差である。一方、系統誤差は、研究の対象者数をいくら増やしても影響を受けずに依然として残る誤差である。系統誤差の存在を十分に認識しておかないと簡単に誤った結論が導き出されてしまう危険性が高い。

〈例1〉高齢男性の平均体重の調査

ごく簡単な例として、ある町の高齢男性の平均体重を調査する研究を考えてみよう。その町は人口100万人で、そのうち高齢男性は10万人いる。そこで、住民票からランダムに抽出した100人の高齢男性を対象に、同一型の体重計を複数用い測定することに決めたとする。これによって、100人の測定値の平均値を用いて、この町の高齢男性の平均体重を推定することが可能になる。この推定値に対する誤差はなにから起こるだろうか。測定値は機器ごとの違い、測定場所、測定を行った時間帯や測定者によって異なるかもしれない。たとえば測定のとときに、どのような場所に体重計を置いていたかは偶然誤差に影響するであろう。この誤差は測定値が高くなることもあれば、低くなることもあるが、誤差を平均すれば高すぎることも低すぎることもないだろう。もしサンプルサイズを100から1,000とか10,000と増やしていけば、このような誤差の影響はほとんどなくなっていく。測定した高齢男性の平均体重と、その町の高齢男性全員の平均体重とを比較してみると、測定のサンプルサイズが大きくなればなるほど誤差がゼロになることがわかっている。しかし、サンプルサ

Naohiro Yonemoto, Mistuhiko Yamada
〒187-8553 東京都小平市小川東町4-1-1

イズを増やしても、影響を受けない誤差もある。たとえば、測定に使用した体重計は、まちがってはじめてからマイナス1 kgに設定されており、すべての体重は、実際より1 kg少なく測定されていたかもしれない。この誤差はどれだけ対象者数を増やしたとしても減らすことはできない。また、ある施設では体重計の少数点を切り上げて計測したが、他の施設では切り下げで計測したとする。このような場合に生じる誤差も、対象者数が増えたからといって減らすことができない系統誤差となる。

偶然誤差は統計学的検定、推定によって対処されるが、系統誤差は研究デザインや交絡の調整などの解析により対処することになる。解析の段階での系統誤差への対応はある程度は可能であるが、実際には「後付け」の対応であり、限界がある。したがって、系統誤差をなるべく少なくするための対応は、研究デザイン（データのサンプリング、測定の方法）計画の段階において十分に検討し、実施することが重要となる。

II. 研究デザインの種類

1. 探索的・仮説生成型研究と検証型研究

研究は目的に応じて、探索的・仮説生成型研究と検証型研究がある。探索的・仮説生成型研究とは、新しい仮説を作り出す、もしくは既存の仮説をより強い仮説へと強化する目的の研究である。症例集積 (case series)、後向き研究 (retrospective study) や前向き研究 (prospective study) といった観察研究はこれに当たる。検証型研究とは、ある仮説に対しての検証を行う実験、介入研究であり、ランダム化比較試験 (randomized controlled trial ; RCT) と呼ばれるデザインがある。それぞれ研究の目的によって研究デザインは異なる。

2. 観察研究

観察研究 (observational study) には、コホート研究 (cohort study)、ケースコントロール研究 (case-control study) などがある。コホートとは

「特定の目的のために選ばれた個人の集団で、ある一定期間追跡されるもの」である。コホート研究は、1つもしくは2つ以上のコホートでの疾患の発症、死亡等の有無を測定する研究で、コホート集団の関心のある曝露 (exposure)、非曝露と疾病の発症、死亡等の有無を比較する。曝露には、喫煙や飲酒、薬剤の投与などの治療もあれば、遺伝子もあるかもしれない。一見単純に思えるコホート研究であるが、だれが追跡対象か、どうやってなにをもって疾病等の発症数を数えるのか、など実際はかなり複雑な準備が必要である。

ケースコントロール研究は、あるコホートのうち、疾患を発症した者と、元のコホートから標本抽出されたコントロールとの比較である。コントロールの抽出方法にはさまざまな方法があるが、抽出は曝露状況と無関係に行われなければならない。ケースコントロール研究にも、コホート内ケースコントロール研究 (nested case-control study)、前向きケースコントロール研究、ケースコホート研究 (case-cohort study)、ケースクロスオーバー研究 (case-crossover study) など、さまざまな研究デザインが開発されており、条件さえあれば効率的に研究を行うことができる。発症のきわめてまれな疾患の研究でケースコントロール研究は有効である。ケースコントロール研究では、直接リスク比 (risk ratio) を計算できないため、オッズ比 (odds ratio) を計算するが、ケースコホート研究ではリスク比を推定することも可能である¹²⁾。

3. 介入研究

介入研究は、コホート研究の一種である。コホート研究では曝露の割り当てはランダムではないが、ランダム化比較試験では曝露をランダムに割り当てることとなる。介入研究は、ある疾患の発症や死亡といった患者の予後指標、アウトカム (outcome) に対して、どの治療法がよいかを検証する場合に用いられる場合が多い。ランダム化比較試験を行う前段階の仮説生成型研究として、過去の成績などのヒストリカルコントロールを用

いた（比較群のない）単群の介入試験を行う場合もある。介入研究は臨床試験（clinical trials）と呼ばれる場合もある。

Ⅲ. 研究デザインの妥当性

研究デザインの妥当性には、内的妥当性と、外的妥当性がある。内的妥当性には、比較の妥当性（もしくは比較可能性〈comparability〉）、対象の妥当性、追跡の妥当性、測定の妥当性、解析の妥当性の5つがある。研究デザインは、まず内的妥当性を高めるための十分な検討をすることが必要である。

1. 比較の妥当性

比較の妥当性とは、比較する2つの要因が比較してよい条件に整っているかどうかという妥当性である。たとえば、2つの治療を比較する場合、同じような患者に2つの別の治療をしているとは考えにくい。通常の診療では、治療の適用は、患者の状態、重症度などに関連している決定しているはずであるので、当然2つの治療を受ける患者の背景は異なっている可能性が高い。このような状況で2群を比較することが妥当であろうか。比較をする場合、比較したい要因以外に、2つの対象者の背景が揃っていなければ妥当な比較であるとはいえないだろう。比較の妥当性とは対照群の設定方法の妥当性でもある。この妥当性が担保されない限り、交絡（confounding）と呼ばれるバイアスから逃れることができない。

交絡の定義とは①リスク要因（結果〈死亡、発症〉と関連している）である、②曝露（治療）と関連している要因である、③曝露（治療）と結果（死亡、発症）の時間関係の間にある変数（中間変数）でない、つまり、その要因が曝露の結果ではない、である¹²⁾。

たとえば、「日常の運動の程度の違いがアルツハイマー病の発症に影響するか」を検討する場合、年齢は交絡因子となるだろう。年齢はアルツハイマー病発症のリスク要因であり、日常の運動の程度とも関連しているだろう。日常の運動の程度に

よって、年齢が進むことはあるはずがない。ほかに、遺伝子、食事や肥満度指数（body mass index; BMI）などは交絡因子であるだろうか。

さらに注意すべき点は、その要因が「いつ」の時点で測定されたかである。曝露よりあとの時点で測定されたものは、直接でなくても間接的に曝露の効果による結果を含むものであるかもしれない。このような間接的に曝露の効果を受けている要因は、交絡因子ではなく、中間変数である。中間変数を交絡として調整してしまうと、かえってバイアスの影響を受けた結果になってしまうので注意が必要である。

ランダム化を行わない観察研究においては、交絡がない、ということは保証されない。よって、観察研究の解析においては、多変量解析などによる交絡の調整の解析を行う必要がある。比較の妥当性を確保する研究デザインが、ランダム化比較試験である。ランダム化という操作によって、比較の妥当性を確保しているのである。

ランダム化とは、比較する要因を対象者に対してランダムに割り当てる方法である。割り当ては乱数を用いて行われる。曜日や月、来院の偶数奇数で割り当てることは、乱数ではないのでランダム化ではない。ランダム化は1つのある集団を同一の特性をもつ2つ以上の群に分割することであり、いわばクローンを作るようなものである。ランダム化されてできた群は、それぞれ測定可能な要因も測定不可能な要因も平均的にリスク要因がすべて均等となり、最も比較の妥当性が高い群を作成することができる。ランダム化が最も比較の妥当性を高める理由は、測定されている要因に関しても、未測定の要因に関しても、平均的に要因が揃えられることである。解析では、未測定の交絡要因に関しては対応できない。ゆえに、ランダム化が強力な方法なのである。

ただし、ランダム化を実施さえすれば交絡がないということを保証するわけではない。たとえランダム化を行ったとしてもサンプルサイズが小さければ、群間の不均衡（性別や年齢がたまたま揃

わないこと) が起こりうる。一方、十分な大きさのサンプルサイズであればその可能性は限りなく低くなる。

交絡の調整を行うデザインには、限定化 (restriction) やマッチング (matching) がある。限定化とは、交絡となる変数が同じ値の人たちだけを対象として限定した研究を実施することである。たとえば、性別で交絡することがすでに明らかなのであれば、女性 (あるいは男性) だけを対象とした研究を実施することである。マッチングは性別や年齢などの条件を揃えてコントロール (比較・対照群) を選定する方法である。マッチングはコホート研究やケースコントロール研究で用いることが可能であるが、それぞれ意味が異なり、解析方法も特別なので、実際にマッチングを用いた研究を実施する場合には十分な注意が必要である¹³⁾。

交絡は、測定されている要因に関しては解析で調整することが可能である。調整の方法には、層別化 (stratification) や回帰モデル (regression model) などがある。近年では傾向スコア法 (propensity score)^{1,13)} や IPW 法 (inverse probability weight) を用いた調整法も用いられる¹³⁾。

2. 対象の妥当性

対象の妥当性とは、研究対象者の適格基準、除外基準と関連している。対象者の選択や研究の参加が系統的に偏ってしまうと、選択バイアス (selection bias) が発生する。選択バイアスがあるかどうかは、研究の対象から外れてしまった者のデータを収集しておかない限り正確に判断することができない。しかし、研究の対象外のデータ収集は実施されない場合が多いため、推測に基づき判断せざるを得ない場合が多い。

また、適格基準や除外基準の定義が不明確であれば、実際の調査対象のなかにそれ以外の者が混ざってしまい、本来調査対象とするべきであった対象者のデータとの間に偏りが生まれることになる。

〈例2〉高齢者のうつ病の調査

ある町における高齢者のうつ病の有病割合を調査する場合を考えてみよう。まず、パイロット試験として、ボランティアとして公募した住民を対象にうつ病のスクリーニングテストを実施した。この調査で得られた結果は、本来調査したかった「その町でのうつ病の有病割合」を反映しているものであろうか。ここで、選択バイアスの影響について注意する必要がある。ボランティアでうつ病の調査に参加した者は、そうでない者に比べて、メンタルヘルスに関心のある者かもしれない。本人もしくは知人、家族にリスクの高い者がいるかもしれない。調査に参加するだけの意欲と体力のある者かもしれない。このような、実際に調査した対象者の特性の偏りは、影響の程度を正確に評価することは困難であるものの、バイアスとして必ず解析結果に影響を与える。こうした選択バイアスは、調査で使用したうつ病のスクリーニング法の性能とはまったく無関係である。選択バイアスへの対策として、スクリーニング法や診断法の開発研究には、ランダム化比較試験を用いることが推奨され、多くの研究で実施されている¹³⁾。

〈例3〉定期検診データを活用した研究

会社等で実施されている定期検診データを活用した研究を実施する場合にも、選択バイアスの影響に注意する必要がある。ある特定の職域を対象とした研究において、その職域における疾患の発症割合と一般集団の発症割合とを比較した報告がなされることがある。しかし、このような比較は期せずして選択バイアスの影響を受けている可能性があるため注意が必要である。ある職業集団を対象としたデータには、疾病の症状や前駆症状のためにすでに退職してしまった者の情報もともと除かれてしまっている可能性が高い。つまり、健康で勤務が可能な者だけが、その職業集団に含まれている可能性が高く、これは健康労働者効果 (healthy worker effect) と呼ばれる。このバイアスの影響を避ける一つの方法として、特定の職業と曝露やリスクが異なる別の職業の集団を比較す

る場合がある。健康労働者効果は次の追跡の妥当性の議論においても注意が必要である。

3. 追跡の妥当性

ランダム化比較試験やコホート研究のように、対象者の予後（アウトカム）を調べる場合には、対象者を追跡することになる。対象者を十分な期間追跡することができればよいが、しばしば脱落（drop out）して、追跡不明となってしまう。とくに、状態の悪い者、もしくはよい者だけが追跡できないとすると、それは研究結果に重大な影響を与える。人年法（person-year method）に基づいた研究のデザインでは、特段の理由によらない打ち切り（censoring）であれば許容されるが、状態が悪化した（もしくはよくなった）から脱落するような、情報を伴った打ち切りであれば、その妥当性は担保できないだろう。またあまりに脱落が多く、十分な期間が追跡できていない場合には、当然、追跡の妥当性に疑問が残るはずである。

4. 測定の妥当性

研究対象の測定についても、バイアスが発症する場合がある。これは、情報バイアス（information bias）と呼ばれる。なかでも、カテゴリ尺度で測定した場合の情報の系統的な誤りを誤分類（misclassified）と呼ぶ。分類のまちがいによって、本来曝露している者がそうでない者に、発症している者がそうでない者になってしまう誤りである。誤分類は発症の仕組みの種類として differential と non-differential に分けられる。疾病の発症や曝露の有無と関連して誤分類を起こしているとすれば、differential であり、そうでないならば、non-differential である。

〈例4〉高齢者の食事とうつ病発症についての調査

たとえば、高齢者を対象に、食事における ω 3脂肪酸の高摂取の有無とうつ病発症について検討する場合を考えてみる。高摂取群か低摂取群かは、摂取量から基準値を設けて、高低の2群に分類することとする。その際に、測定誤差が生じて、どちらかに誤って分類されることが起こりうる。本

当は高摂取でないが、低摂取群になる場合もあれば、低摂取にもかかわらず、高摂取群に分類されることもある。この誤分類がうつ病の発症と関連していなければ、これは non-differential である。non-differential な誤分類は、多くの場合、効果を薄める、つまり、差がないという結果に導く方向にバイアスがはいることが知られている^{12,13)}。

ケースコントロール研究などの後向き研究で起こるのが、思い出しバイアス（recall bias）と呼ばれるものである。これは differential な誤分類である。すでに疾病が発症したあとに、インタビュー調査などで過去の曝露情報を調査する場合に起こる。「疾患を発症した者」は、発症原因や思い当たる背景、曝露について詳しく思い出し、それが疾患の発症と関連して考えるかもしれない。ところが、コントロールである「その疾患が発症していない者」は、過去の曝露状況をそれほど詳しく思い出せないかもしれない。このように思い出し方に違いが生まれるだろうことは容易に想像できる。思い出しバイアスをできるかぎり防ぐために、正確に偏りなく記憶をたどれるように質問法を工夫する、面接調査だけでなく、過去に記録された診療録など別の情報源からも情報から得るという方法もある。このような differential な誤分類は効果を過大評価することも、過小評価することもありうる^{12,13)}。

5. 解析の妥当性

どのような研究であろうと、収集されたデータについて研究デザインにあった解析が行わなければ、妥当な結果を得ることはできない。データの種類、分布等のデータの性質、研究デザインによって解析方法は異なる。たとえば、平均値で比較することが妥当でないデータに対して、平均値の差の比較の解析を行ったとしても、その結果は不適切である。研究デザインと解析方法は常にセットであり、事前に研究計画書に記載するべきである。また、解析も解析計画書を作成してから実施する。近年では、とくに介入研究では、後付けの解析や不都合なデータの隠蔽を防ぐため、臨床試

験登録により、あらかじめ解析するアウトカムを提示する義務があり、また論文投稿の際に、解析計画書の提出を要求する雑誌までみられるようになってきた。収集したデータを事後的に処理し、都合のよい結果のみを創出しようとすることは慎むべきである。

6. 外的妥当性、一般化可能性

一方、内的妥当性が高まったとしても、その結果が現実の臨床場面に適用できなければ意味がない。この妥当性を、外部妥当性、一般化可能性 (generalizability) と呼ぶ。とくに治験等の介入試験では、日常臨床としては適用となるであろう対象者が、リスクが高いため、研究のレベルでは除外される場合も少なくない。うつ病研究でも、希死念慮の高い者を対象者から除外して実施されることが多い。また、研究結果を、ある程度似た他の集団に外挿する必要がある場合もあるであろう。その場合には、研究結果の解釈には十分な注意が必要である。このような研究結果と日常臨床との間にギャップを埋めるため、次のステップとして、日常臨床に近いかたちで実施する実務的 (pragmatic) な試験を実施する必要がある。

7. 研究デザインの種類とエビデンスレベル

研究デザインの種類によって、得られる研究結果の妥当性は異なっている。EBM (evidence based medicine) においてよく語られ、アメリカ AHRQ (Agency for Healthcare Research and Quality) や Oxford Centre for Evidence-based Medicine などが提示しているエビデンス (evidence) のレベルとはこれを指しており、多くの場合、バイアスがある可能性の程度をもって示されている。ただし、小規模なランダム化比較試験よりも、適切にデザインされた前向き数千例の大規模なコホート研究のほうがむしろ真実に近い結果を反映している場合もありうる。そのため、「研究の質」をエビデンスのレベルのみをもってして判断することには限界があり、個々の研究で研究の妥当性を評価する必要がある。

IV. 継時データの研究とデータの欠測

老年精神医学の領域でよく用いられる研究デザインに、継時データの研究がある。これはコホート研究、もしくはランダム化比較試験において、時間経過に伴う変化を評価する研究デザインである。継時データの研究では、個人1人ひとりが、決められた時点ごとに、繰り返しデータ測定 (repeated measures) される。臨床家にとっては、ある期間の患者の健康状態の推移を追跡することが必要である。また、患者に対する治療がたった一度で終了することはまれである。よって、継時データの研究におけるアウトカムに関する情報は、研究開始前および研究期間中のいくつかの時点で測定される必要がある。このような継時データを用いた研究は、精神医学研究の多くの領域で、とくに重要なデザインとなっている³⁾。

しかし、このような研究デザインでは、データの欠測が生じやすい。ある対象者が脱落すればそれ以降のデータはいっさい入手できない。また、すべての対象者を決まった時点で測定できない場合もある。ある1時点のデータが1つでも欠測しているからといって安易に解析対象から除いてしまうことは、対象者数を減らし、さらにバイアスを発生させることになる。たとえば、抗うつ薬の臨床試験の脱落の割合は3か月の治療期間で30~40%以上にもなるといわれている³⁾。脱落や欠測値の数が少なかったとしても、予後の最も悪い者ばかりのデータが脱落・欠測しているとか、逆に予後がよい者ばかり脱落・欠測していれば、最終的な解析結果に大きな影響があるであろう。そのため、欠測値を含むデータの取り扱いには十分な注意が必要である^{2,3,5,6,8-10)}。Rubinは、データの欠測を、その発生の仕方によって、完全にランダムな欠測 (MCAR)、ランダムな欠測 (MAR)、無視できない欠測 (non-ignorable missing) に分類した^{3,6)}。患者の脱落やデータの欠測が及ぼす影響については、この分類を基に考えると理解しやすい。

1. 完全にランダムな欠測

完全にランダムな欠測 (missing complete at random; MCAR) は、以前に観察した値にも、欠測の値にも、どちらの値にも依存しない確率で起こる欠測である。結果として、観察された(欠測でない)値は、全体の対象者の単純なランダムサンプルと同じことになる。ありそうな例としては、実験室での測定で、試験管を落としたために測定できなくなったとか、研究対象者の偶然の事故死、もしくは、別の地域への異動などがある。継時データにおいて、部分的に起こる欠測値、たとえば、患者が一時的な理由(たとえば、遊びに行ったため)により来院しなかった場合などは、完全にランダムな欠測を考えるのに理にかなった仮定である。この種の欠測は結果にほとんど影響を与えないが、本当にランダムな欠測であるかはどうか、かなり強い仮定である。

2. ランダムな欠測

ランダムな欠測 (missing at random; MAR) とは、欠測が過去に観測されていたアウトカムに依存した確率で起こる種類の欠測である。ここでの欠測は観察されたデータだけに依存する。もし同じ背景要因をもち、特定の時点まで、またそれを含む時点で、アウトカムが同じであるなら、その時間に残っている対象者の特徴の分布と、特定の時点で脱落した者の特徴の分布は同じはずである。このタイプの欠測の例として、アウトカムが収縮期血圧である降圧薬の例がある。この研究では、対象者の血圧が高くなりすぎたとき、研究から除外するとしている。参加者が脱落するまえに、脱落時点の血圧が観察されている。これは血圧の値に依存しているので、欠測のメカニズムは完全にランダムな欠測ではないが、ランダムな欠測である。観察されたデータの一部だけに依存して欠測しているからである³⁾。もう1つ、BMIを測定した研究の例を紹介する。BMI測定値の欠測は、介入期間に体重が増えたか減ったかどうかにかかわらず、早期に来院した者でBMIが高い者は、面倒くさがって、あとの来院時に測定をしたがらな

かった、と考えたとする。この場合、その欠測はランダムな欠測であるが、完全にランダムな欠測ではない³⁾。

3. 無視できない欠測

最後の無視できない欠測 (non-ignorable missing, しばしば情報をもった欠測といわれる) は、測定していない値に依存した確率で起こる種類の欠測である。これは、患者から観察されたアウトカムが、系統的に通常より高かったり低かったりする場合において、観察データが欠測する場合を指す。現実的には、おそらく対象者の状態が悪くなったり、よくなったりしていることに関連しているであろう。たとえば、低所得者もしくは超高所得者インタビューで個人の収入に関して回答したがるらない、ということがある。ほかにありそうな例は、対象者の血圧があまりにも高くなったりすると、この値は観察されなくなったり、耐えられないくらいの痛みになると、関連する痛みの値は記録されていなかった、などがある。上で紹介したBMIの例では、最後の外来から体重がさらに増えていたら、対象者が測定をより避けがちになるとすれば、そのデータの欠測は無視できない欠測となる。この種類の欠測値が含まれているデータは扱うことはきわめてむずかしい³⁾。

欠測値の補填に関しては、古くから欠測値の最終観察値で補填するLOCF法 (last observation carried forward) がよく用いられてきたが、近年、これは推奨されるべき方法ではないと指摘されている⁷⁾。データ欠測の分類等によって、混合効果モデル (mixed effect models) や多重補填法 (multiple imputation)、感度解析 (sensitive analysis) などの解析方法が開発されており、臨床論文でも多く用いられるようになってきている^{3,7,10,13)}。

おわりに

老年精神医学研究にかかわらず、どのような臨床研究にも、その妥当性を損なわせるさまざまなバイアスが入り込む余地がある。そして、このよう