

原 著

模擬試験成績とヒューマンエラー発生数との関係について

獨協医科大学 国試教育センター

一杉 正仁 菅谷 仁 平林 秀樹 妹尾 正
上田 秀一 下田 和孝 田所 望 古田 裕明

要 旨 試験におけるエラーの発生予防策を立案して実践する上では、エラーを頻回におかす学生を早期に発見して、重点的に介入する必要がある。そこで、学生の試験成績とヒューマンエラー発生との関係を明らかにした。対象は医学部6年生102人が受験した医師国家試験模擬試験結果である。自己採点結果と、マークシートによる機械的採点結果をもとに予期せぬ失点の原因を調べた。また、試験成績とエラー発生頻度との関係を調べた。ヒューマンエラー発生の有無によって、試験成績に有意な差はなかった。しかし、400問あたり3問以上のエラーをおかす受験生の得点率は $61.4 \pm 6.1\%$ であり、2問以下のエラーをおかす受験生の得点率($65.6 \pm 5.3\%$)に比べて有意に低かった。成績上位者を含めた全ての受験生にエラーが生じ得るため、全学年を対象とした注意喚起が必要と思われる。また、成績下位者には多数のエラーをおかす学生が多くみられることから、これらの学生を対象とした重点的な教育介入が求められる。

Key Words : 試験問題, 成績, ヒューマンエラー, 医学教育, 医師国家試験

はじめに

医師国家試験は3日間に500問の多肢選択問題(MCQ)に解答することで、その医学的知識が確認される。われわれは医師国家試験模擬試験において、受験者がおかす予期せぬ失点の原因を調べてきた^{1,2)}。そして、問題用紙で選択した解答肢をマークシートに記入する際に誤ること、マークシートに記入する以前に、問題文末に指示されている選択肢数を誤認すること、が主たる原因であることがわかった。これらすべてのエラーを総括して、試験におけるヒューマンエラーと呼ぶ。これら全てのエラーは認知システム工学分野で、「望ましからざる結果が生じ、その原因が人の身体的あるいは心的行為の何らかの側面にかかわるような状況」と定義されている³⁾。そこで、試験におけるヒューマンエラーの実態を分析することで、受験生の認知・判断・行動特性を明らかにし、予防対策を講じることができると考えた。

これまで、ヒューマンエラーの実態が解明され予防策が検討されてきたが^{4,5)}、受験者の成績とエラーの発生

頻度については明らかにされていない。今後、エラーの発生予防策を立案して実践する上では、エラーをおかす頻度が高い学生を早期に発見して、重点的に介入する必要がある。そこで、学生の試験成績とヒューマンエラー発生との関係を明らかにし、今後の介入教育のために有用な知見を得たので報告する。

対象および方法

対 象

102人の医学部6年生が2009年12月にTECOM 104-3の医師国家試験模擬試験を受験した。受験生は1ヶ所に集まり、縦型のマークシートを用いて受験した。試験問題は5つの選択肢から1つの正解肢を選択するAタイプ、2つの正解肢を選択するX2タイプ、3つの正解肢を選択するX3タイプ、多肢から1つの正解肢を選択するLタイプのMCQ 500問で構成されている。試験問題はA~Iの9つのカテゴリーから構成され、各論、総論、必修領域に3カテゴリーずつ配置されている。それぞれの領域で、単純な知識の想起や与えられた情報を理解して解答する一般問題と実地臨床に関する文章を理解して具体的な問題解決を求める臨床問題が含まれ、全体として、それらは250問ずつとなっている。なお、カテゴリーC, F, Hは必修領域の問題であるが、全ての問題が択一問題で構成されており、エラーがほとんど発

平成22年12月13日受付, 平成23年1月6日受理
別刷請求先: 一杉正仁

〒321-0293 栃木県下都賀郡壬生町北小林880
獨協医科大学 法医学講座准教授

表1 試験カテゴリー別にみた、ヒューマンエラーの分類別一覧

| | カテゴリー | | | | | | 計 |
|------------------------|-------|------|------|------|------|------|-----------|
| | A | B | D | E | G | I | |
| 問題数 | 60 | 62 | 60 | 69 | 69 | 80 | 400 |
| 複択問題の割合 (%) | 38.3 | 37.1 | 36.7 | 40.6 | 27.5 | 47.5 | 30.6 |
| エラーの分類 | | | | | | | 計 (%) |
| タイプ1 問題を解答する際のエラー | | | | | | | |
| 1-A 指示より少ない肢を選択した | 11 | 7 | 8 | 9 | 1 | 9 | 45 (41.3) |
| 1-B 指示より多くの肢を選択した | 5 | 6 | 2 | 2 | 4 | 9 | 28 (25.7) |
| 1-C 問題を解き忘れた | 0 | 0 | 0 | 1 | 0 | 0 | 1 (0.9) |
| タイプ2 マークシート記載時のエラー | | | | | | | |
| 2-A 選択したものと異なる記号をマークした | 0 | 3 | 1 | 5 | 1 | 7 | 17 (15.6) |
| 2-B マーク数を誤った | 1 | 1 | 3 | 3 | 3 | 3 | 14 (12.8) |
| 2-C マークしなかった (ノーマーク) | 1 | 1 | 0 | 1 | 1 | 0 | 4 (3.7) |
| エラーの総数 | 18 | 18 | 14 | 21 | 10 | 28 | 109 (100) |

生しないことが指摘されている²⁾。さらに、当該模擬試験においては、必修領域はほかの領域と区別されて採点が行われることから、本検討では医学総論および各論の範囲である400問を対象に解析を行った。それぞれのカテゴリーにおける複択問題の割合（総問題数に対してX2タイプとX3タイプの問題がしめる割合）を表1に示す。なお、試験時間であるが、医師国家試験と同様に、一般問題で1問あたり1分、臨床問題で1問あたり2.5分を目安に設定された。

方法

受験生が問題用紙に記入した正解肢をもとに集計した自己採点内容と、マークシートをもとに機械的採点を行った内容を比較した。そして、もれなく全ての問題に解答していたか、解答時に問題の指示文に対して正確に対応していたか（単肢選択や多肢選択が正しく行われていたか）、問題用紙上で選択された解答肢が正しくマークシートに転記されていたか、について問題ごとに調査した。さらに、全てのエラーを既報のように分類した（表1）^{4,5)}。

得点率であるが、医師国家試験に準拠して、一般問題は1問1点、臨床問題は1問3点として採点された。対象問題は一般問題と臨床問題それぞれ200問ずつで構成されているので合計800点となる。個々の受験生について、その得点率（素点/800点）を算出した。なお、平均得点率は平均±標準偏差で記載した。

結果

概略

全ての受験生は時間内に試験を終了していた。対象者

の平均得点率は $64.9 \pm 6.5\%$ であり、 $46.9\% \sim 75.6\%$ に分布していた。受験生がおかした全てのエラーを表1に示す。総エラー数は109であり、受験生1人が400問の試験問題に対して、平均1.1個のエラーをおかしていた。

次に、対象受験者一人あたりがおかすエラーの頻度を調べた。すなわち、それぞれの試験における受験者に対して、400問あたり何問のエラーがあったかを調べた。まったくエラーをおかしていない受験者は半数以下の48人（47.1%）であり、エラーを1問おかした受験者は23人（22.5%）、2問では14人（13.7%）、3問では11人（10.8%）、4問では5人（4.9%）、5問では1人（1.0%）であった。

エラーの内容と発生頻度

エラーのタイプ別では、試験問題を解答する際にエラーをおかす1型が67.9%と多くを占めた。詳細な分類では、2肢選択すべきところを1肢しか選択しなかったというような、指示より少ない肢を選択した1A型が41.3%と最も多く、1肢選択すべきところを2肢以上選択したというような、指示より多くの肢を選択した1B型が25.7%と続き、この両方で1型エラーのほとんどを占めた。

エラーの発生数と得点率との関係

まず、エラーの有無別に得点率を比較した。すなわち、エラーをおかさなかった48人と、1問以上のエラーをおかした54人との間で平均得点率を比較した。その結果、前者で $65.6 \pm 5.2\%$ 、後者で $64.3 \pm 6.0\%$ と統計学的有意差を認めなかった（unpaired t-test, $p > 0.05$, 図1）。次に、エラーの発生数ごとに、対象者の平均得点率を比

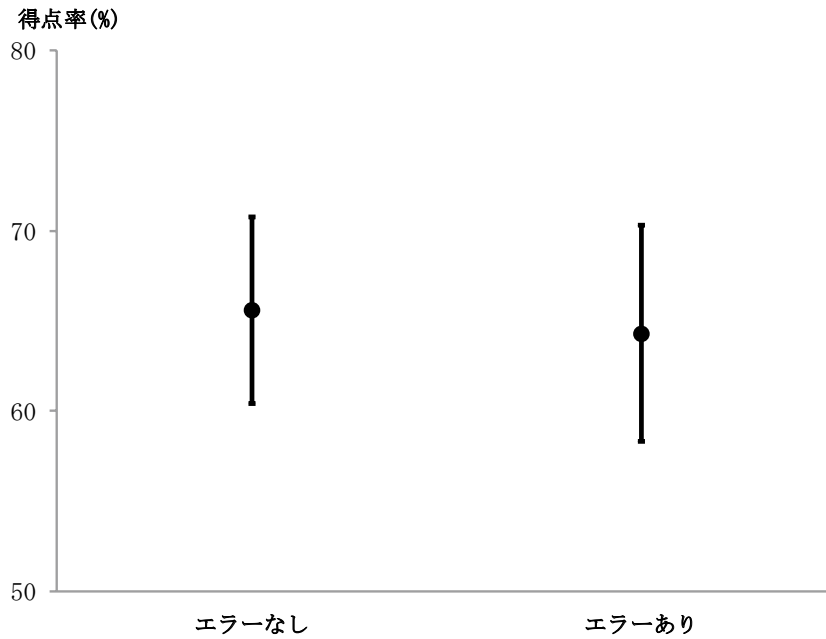


図1 ヒューマンエラーの有無と試験成績との関係

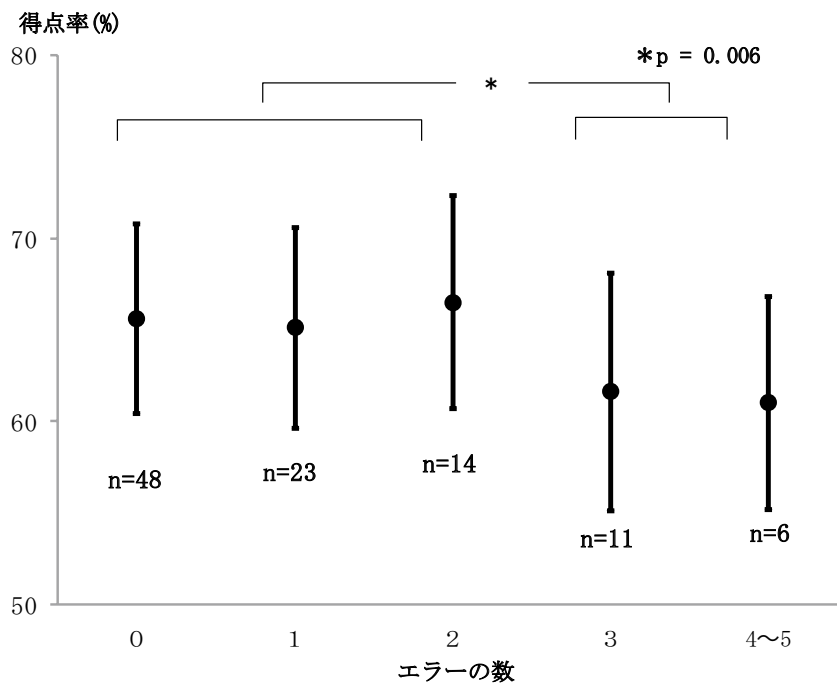


図2 おかしたヒューマンエラーの数と試験成績との関係

較した (図2)。エラーの数が0, 1, 2, 3, 4~5の5群で平均得点率を比較したが, 統計学的有意差は見られなかった (Kruskal-Wallis test, $p > 0.05$)。しかし, エラーの数が0, 1および2の3群ではいずれも平均得点率が65.1%以上であるが, エラーの数が3および4~5の2群では61.6%以下であった。そこで, エラー数が2以下 (85人) と3以上 (17人) の2群に分けて, 対象者の平均得点率を比較した。その結果, 前者で 65.6 ± 5.3

%と後者の $61.4 \pm 6.1\%$ に比べて有意に高かった (Mann-Whitney test, $p = 0.006$)。なお, エラーによる得点率の低下を考慮し, 全てのエラーがなかったと仮定した場合の得点率を算出した。すなわち, エラーをおかした受験者の得点にエラーによる失点分を加えて, 再度得点率を算出した。その結果, 前者で $65.7 \pm 5.3\%$, 後者で $61.9 \pm 6.1\%$ と換算前とほぼ同様の値であり, 統計学的有意差を認めた (Mann-Whitney test, $p = 0.01$)。

考 察

人間は日常生活で様々な失敗をおかすことがある。特に、決められた時間内に500問のMCQに解答するという条件下で、長時間の思考と緊張状態が加われば、エラーが発生しやすくなると考えられる。われわれは、この実態を把握し、その予防に努めてきた。

対象者の約半数が1問以上のエラーをおかしていた。そして、エラーの種類では、問題文の指示よりも少ない数の正答肢しか選択していないものが最も多かった。これは、既報の結果と同様の傾向である⁴⁾。また、表1では、複択問題の割合が多くなる程、エラーの発生数が増加する傾向であったが、これも既報と同様の傾向である^{2,4)}。すなわち、単肢選択問題と多肢選択問題が混在する試験では、受験者は問題文を最後まで注意深く読まなければならない。したがって、終始高い集中力を維持することが要求される。このようなタスクが加わることで、エラーが誘発されやすい環境になったと考えられる。

エラーの発生率と成績との関係であるが、エラーの有無で比較すると得点率に有意差はなかった。したがって、成績上位者を含めたすべての学生にエラーが起こり得るということを支持する結果であった。われわれは、教育的介入によって、エラーの発生数を低減できることを明らかにしたが、エラーを完全になくすことは困難であることも指摘してきた^{3,4)}。今回の結果は、これを示唆するものである。したがって、3日間で500問のMCQに解答するという状況では、必然的にヒューマンエラーが生じると考えて良いだろう。

1問あるいは2問のエラーをおかした人の得点率は、エラーをおかさなかった人の得点率とほぼ同様であったが、3問以上のエラーをおかす人で、得点率は有意に低下した。多くのエラーをおかす受験生は成績下位者に多いといえる。成績下位者では正しく問題を理解し、一定の思考のもとに解答を得て、それを適切に記載するという一連の処理を行ううえで、集中力が欠けるのであろう。したがって、まず、これらの学生を対象に重点的な教育的介入を行うべきと思われる。

うっかりミスと言われているタイプのエラーは、機械などの偶発的故障と同様に、確率統計的に起こると考えられている。機械の信頼度は故障回数が少ないことで評価されるが、人間行動の信頼性はエラーの発生頻度が少

ないことでも表される³⁾。エラーが発生する背景にはさまざまな原因があるが、人間の問題と周囲の問題に分けて考えられる⁶⁾。試験という環境は変えられないため、今回みられたエラーを予防するには、人間の工夫が必要となる。人間は注意が散漫になると、必要な情報を見落とし、エラーを起こしやすい。したがって、試験中に終始注意力を保つ訓練が必要である。多くのヒューマンエラーでは、エラーをおかした本人がそれに気づかないことが多いという⁷⁾。エラーの実態を明らかにし、これを受験者本人にフィードバックすることは、学生自らが問題点に気付き、そして再発を予防するための行動変容を遂げるうえでも重要である。したがって、今後も全学生を対象に、エラーの実態を調査する必要がある。

本検討では、エラー予防の重点教育を行う対象が明らかになった。さらに、全ての学生を対象に、エラー予防の注意喚起を行う必要があることも示唆された。

本検討結果の一部は第42回日本医学教育学会大会(平成22年7月、東京)で発表した。

文 献

- 1) 一杉正仁, 菅谷 仁, 平林秀樹, 他: 医師国家試験模擬試験におけるマークミスの解析. *Dokkyo J Med Sci* **35**: 79-84, 2008.
- 2) 一杉正仁, 菅谷 仁, 平林秀樹, 他: マークミスの発生率が高い試験問題とは. *Dokkyo J Med Sci* **35**: 175-178, 2008.
- 3) 古田一雄: 認知システム工学と安全. 医療安全管理事典. 長谷川敏彦(編), 朝倉書店, 東京, pp12-16.
- 4) 一杉正仁, 菅谷 仁, 平林秀樹, 他: 医師国家試験模擬試験におけるヒューマンエラーの実態と教育効果について. *医学教育* **41**: 119-124, 2010.
- 5) 一杉正仁, 菅谷 仁, 平林秀樹, 他: 試験におけるヒューマンエラーの予防対策について. *Dokkyo J Med Sci* **37**: 77-81, 2010.
- 6) 日本航空技術研究所: ヒューマン・ファクターガイドブック, 日本航空技術研究所, 東京, p.1-139, 1995.
- 7) 松尾太加志: 心理学と安全. 医療安全管理事典. 長谷川敏彦(編), 朝倉書店, 東京, pp17-21.

Relationship between the occurrence of inadvertent Human Errors and Scores in the Trial Examination

Masahito Hitosugi, Hitoshi Sugaya, Hideki Hirabayashi, Tadashi Senoo, Shuichi Ueda,
Kazutaka Shimoda, Nozomu Tadokoro, Hiroaki Furuta

Medical Education Center for National Examination, Dokkyo Medical University School of Medicine

To examine the relationship between the occurrence of inadvertent human errors and examination scores, the results of the trial examination were analyzed. The 102 of 6th grade medical students took the trial examination for the National Examination for Physicians' which consist of 500 multiple-choice questions. More than half of the students (52.9%) made inadvertent errors during the examination. Two-third of errors were either selecting the wrong number of answer options (i.e. when a two choice selection was required, only one choice was selected).

There was no significant relationship between the occurrence of inadvertent human errors and their scores. How-

ever, the mean score of the students who made more than two errors in 400 questions was $61.4 \pm 6.1\%$, significantly lower than those of other students ($65.6 \pm 5.3\%$). Attention to avoid the inadvertent human errors during the examination might be emphasized for all students. Furthermore, the intervention to reduce the inadvertent human errors is especially needed for the students with low scores.

Key words : examination, inadvertent human error, score, medical education, National Examination for Physicians'