

108. インターネットを活用した医学教育評価システム (Web-QME) の開発と検証

福原俊一¹ 松村真司² 木内貴弘³ 日合 弘⁴
加我君孝⁵

(¹京大・医学研究科健康解析学, ²東大・医学教育国際協力研究センター, ³大学病院医療情報ネットワーク (UMIN), ⁴京大・医学研究科病態生物医学, ⁵松村医院)

医学教育の学生による評価は、医学教育の質の改善のために重要な役割を果たす。しかし評価結果の収集・集計・解析の煩雑さや、教官のプライバシーの保護の不完全性、教官へのフィードバックが円滑に行われない、などの理由で、実施が困難なことや初期の目的が達成できないことが多い。また全国のどの医学教育機関でも使用可能な標準化したシステムも存在していない。目的：インターネットを活用した医学教育評価システム (Web-QME) の開発と実施可能性の検討。方法：大学病院医療情報ネットワーク (UMIN) のサーバーを活用して、学生が簡単にアクセスでき授業評価を入力でき、結果を瞬時に集計・図表化して教官や科目責任者にフィードバックが可能、かつデータのダウンロードも可能なシステムを開発した。さらにこのシステムの実施可能性を施設数、教科で検討した。結果：企画、内容作成、プログラミング、パイロットスタディなどの試行改訂の作業のために約1年半を要した。パイロットスタディでは学生からのアクセスや入力の問題がなかったが、フォーマットなどの改訂を促された。また結果集計、図表化、ダウンロードなどは、未経験者でも容易に実施可能であった。結論：インターネットを活用した医学教育評価システム (Web-QME) バージョン1.0が完成し、実施可能性が確認された。今後は施設数と科目数を増やして、実証実験を継続する。また、評価項目も、現在の「コア項目」に加え、実習や臨床研修に通した内容のものを「オプション項目」として追加する予定である。

109. 医師国家試験への Computer Adapted Test 導入の試み

高林克日己¹ 鈴木隆弘¹ 池田 央²
福井次矢³ 植村研一⁴ 細田達一⁵

(¹千葉大・医学部附属病院医療情報部, ²(株)教育測定研究所, ³京大・医学部総合診療部, ⁴弘慈会加藤病院, ⁵厚原記念病院)

背景：CBT が共用試験として開始されたが、医師国家試験についても米国の Step 3 に相当したコンピュータ化試験の導入が論議されている。米国ではさらに問題の正答誤答で次問題を決定していく Computer Adapted Test (CAT) が看護師の認定試験などに導入されている。目的：医師国家試験への CAT 導入の是非・可能性について検討する。方法：すでに施行されて正答率、識別指数の分かっている 800 問の問題を利用し、正解率の高い受験生と低い受験生を想定してシミュレーションを 100 問行い、どの程度の判定が可能かを検討した。また正答率のみで作成した CAT で実際に学生に試験を実施し、アンケート調査を行った。結果：100 問のシミュレーションでは正解率の高い受験生も低い受験生も安定した判定域には入らなかった。実際に行った学生のアンケートでは 67% が紙のテストより大変だったと答え、前問に戻れないこと、メモを記入できないこと、集中して疲労度が大きいことを理由に挙げた。長所としては選択が容易 (マークを塗らなくてよい)、画像がきれいである、ことなどが挙げられた。CAT に対しては 22% が不公平さを、22% が異和感を指摘した。全体としては 63% が従来の紙のテストを希望した。考察：CAT の導入により各問の解答ごとに受験生のレベルを評価し次問題を選択することで、従来より少ない問題数での判定が期待できる。しかし今回のトライアルでは問題数が少ないために十分に適切な次問題が用意できず、少量の問題数での判定は困難であった。また学生のアンケートでは紙の試験をより好ましいとする意見の方が多かった。今回は CAT のためというより準備不足のために思わしい結果でなかったことも考えられ、これらを含め今後 CBT, CAT の医師国家試験の導入を考えるときに解消されなければならない問題点を検討する。

* *

要望演題

共用試験を巡る問題：CBT と OSCE (口演)

110. 12 ステーション OSCE に関する評価者、学生、SP へのアンケート調査

上野隆登 吉田一郎 犬塚裕樹 佐田通夫
(久留米大・医学部基本的臨床技能実習委員会)

目的：久留米大学医学部では医学科 4 年生を対象とした本年度の OSCE で 2 校の医科大学から外部評価者を迎え「医学教育モデル・コア・カリキュラム」に対応した 8 課題を包括する 12 課題の共用試験 OSCE を実施した。OSCE 実施後評価者、学生、SP へのアンケート調査を行い興味ある結果を得たので報告する。方法：医学科 4 年生を対象に「医学教育モデル・コア・カリキュラム」に対応した医療面接、頭頸部・胸部・腹部・神経の各診察、脈拍・血圧測定、救急処置、外科基本手技の 8 課題と、カルテ記載、画像診断、心音聴取、クロスマッチを含む 12 課題の共用試験 OSCE を実施した。OSCE 実施後評価者、学生、SP に対するアンケート調査を行った。結果：1) 評価者；ステーションの数についてちょうど良い 70%、疲労を感じない 68% であった。課題内容は満足 75%、評価の均一性では均一にできた 77%、試験時間がちょうど良い

78%、フィードバックの時間と必要性ではおのおのちょうど良い 64%、必要 89% であった。2) 学生：OSCE 終了後の疲労感では疲れた 94%、理想のステーション数は 10~15 ステーションが 48%、試験への満足度では 64% が不十分、客観的に評価されたと感じた学生は 65% であった。またステーションにより希望する試験時間、フィードバックの有用性は異なっていた。3) SP；12 名の学生の模擬患者役について疲れを感じない 75%、均一に対応できた 100%、連続対応できる学生数は 15~30 人が 75%、5 分の試験時間はちょうど良い 100%、また身体診察も引き受けるとの回答者 100% であった。結論：1 課題 5 分、フィードバック 1.5 分、12 ステーション OSCE は評価者、学生、SP にとっておおむね好評であった。課題内容、途中の休憩の必要性、SP の身体診察の可能性などの問題点の検討が共用試験医科 OSCE の実施に向け必要と思われる。

Medical Education (Japan)

医学教育 第34巻・補冊 (通巻第202号)

平成15年7月25日発行

編集・発行 日本医学教育学会

〒113-8622 東京都文京区本駒込5-16-9

(財)日本学会事務センター内

電話 03-5814-5801 FAX 03-5814-5820

発 売 株式会社篠原出版新社

〒113-0034 東京都文京区湯島2-4-9 MDビル

電話 03-3816-5311 FAX 03-3816-5314

定価1部1,680円(本体1,600円)送料85円

〔日本医学教育学会に入会希望の方は上記学会事務所に
に葉書または電話でご連絡下さい。〕

本誌の予約(会員外の方)について

本誌の予約購読引受期間は2~12月の1年間に限りま
す

(途中月からの購読をご希望の場合は当月から12月まで
の誌代をお送り願います)

本誌は最寄りの書店または発売元あてに送り先を明記
し、送料を加算した前金でお申し込み下さい。

特別号・臨時増刊号などの定価はそのつど決めますの
で、別に精算をお願い致します。

発売元直接のご予約についての誌代切れご案内は、誌
代の切れる前月にお送り致します。誌代が切れますと送
本を中止しますので、お早めにお払い込み下さい。

送本先変更の場合は、新旧のあて先に雑誌名を併記し
てお知らせ下さい。

教育評価への新しい試み

—教育者≠評価者による学生評価—

福島 統 (東京慈恵会医科大学 医学教育研究室教授)

財団法人 大学基準協会 大学評価・研究部

『大学評価研究』第4号抜刷

(2005年2月発行)

教育評価への新しい試み

—教育者 ≠ 評価者による学生評価—

福 島 統

東京慈恵会医科大学
医学教育研究室教授

はじめに

1. スクール・ミッションを具体化するカリキュラム構造の必要性
2. 学生評価へのあり方への疑問
3. 東京慈恵会医科大学総合試験システム
4. 評価を大学が管理すること
5. 評価とは学生へのメッセージ

おわりに

[キーワード]

評価、総合試験、構造化されたカリキュラム、Capability、内省的実践

はじめに

東京慈恵会医科大学では、1996年から従来の伝統的カリキュラム（講座が教育を担当）から統合カリキュラムに変更した。統合カリキュラム導入に伴い、今までの講座（科目）別試験を全廃し、統合カリキュラムを構成するコース（16単位～30単位でまとめられた科目群）ごとに学生評価のための「総合試験」を行っている。本論文では、東京慈恵会医科大学が行っている「総合試験」の考え方、特に、評価を大学が管理することで大学が学生に提供している教育の質をどのように保証しているのかについて述べる。東京慈恵会医科大学は医学部単科大学ではあるが、東京慈恵会医科大学のこの試みは、専門職業職者養成を使命にする大学の参考事例になると考えている。なお、この試みは平成15年度の「特色ある大学教育支援プログラム」に採択された。

1. スクール・ミッションを具体化するカリキュラム構造の必要性

1995年までは、東京慈恵会医科大学では伝統的カリキュラムが採用されていた。すなわち、2年間の教養教育の後、4年間の専門教育があり、専門教育課程は、解剖学、生理学、生化学、病理学、内科学、外科学などの講座による教育が行われていた。各科目を担当する講座は、独自にその科目の教育目標を設定し、学生教育を行っていた。さらに、同じ解剖学でも、第1講座と第2講座があり、解剖学第1講座は肉眼解剖学と神経解剖学、第2講座は組織学と発生学を担当していたが、それぞれの教育内容の調整はされておらず、講座中心の学生教育が行われていた。このように、各講座が独自に教育目標を設定し、学生評価での合格ラインも各講座が独自に決定していたため、Structured Curriculum（構造化されたカリキュラム）にはなっていなかった。大学には、特に私立大学には建学の精神というスクール・ミッションが存在する（東京慈恵会医科大学のスクール・ミッションは「病気を診ずして、病人を診よ」である）。しかしながら、講座中心の教育では、その講座の主任教授一人の考えで科目の教育目標が設定されるため、各科目の教育目標の総和が大学の教育目標（スクール・ミッション）にはなっていなかった。さらに、教育が講座によって支配されていたため、大学が各科目の教育に十分には介入できず、大学としての学習目標（スクール・ミッション）がカリキュラムに具体化できる体制とはなっていなかった。

1996年、今までの反省に立ち、東京慈恵会医科大

学では6年一貫教育を統合カリキュラムに改組した。いままでの講座による教育科目を全廃し、新たにコース・ユニットに組み替え、完全にカリキュラムの中から講座制をなくした。コースを、総合教育、人体の正常構造・機能、病因・病態、臨床医学、臨床実習、選択実習、医学総論などの大きなテーマごとに設定し、各コースにはユニットと称する教育単位を組織し、カリキュラム構造を明確化した。2年生後期のコース「基礎医科学Ⅱ」のテーマは臓器別正常構造・機能であり、このコースには、循環器系、呼吸器系、神経系、生殖器系などのユニットが配附されている。コースには教育責任者としてコース責任者が決められ、コース責任者が各ユニットの責任者を指名する。ユニット責任者は学内外の教員を自由に指名し、ユニットの教育計画を実施する。ユニット責任者には教育予算執行権もある。東京慈恵会医科大学では、ユニット責任者からの講義・実習依頼は、業務命令と定義されている。従って、ユニット責任者は講座の枠を超えて、カリキュラム遂行のために誰にでも教育業務命令を出すことができる。余談ではあるが、私が解剖学の講師の時代にユニット「運動器系」のユニット責任者になったときに、当時の整形外科講座の主任教授に何月何日何時間目の2年生に「関節障害」の講義を命ず、という音類を出したことがある。もちろん、講義をしていた。

コース責任者とユニット責任者は教育に関して絶大な権限を持つため、その任期は1年間である。このように講座とは別に教育責任体制としてのコース・ユニット制を確立した。

カリキュラム改革の時、大学の教育目標（スクール・ミッション）の確認を行い、その目標が卒業時に全学生に達成されるための各コースの教育目標を設定し、そのコースの目標に沿って各ユニットの教育目標を設定しなおした。学生はコースを終了することで、確実に次のコースを受講するレディネスを整え、次のコースで十分な学習ができるよう、そして卒業時には大学が設定したスクール・ミッションに近づくことができるStructured Curriculumを作った。

医科大学では、従来、講座には教育、診療、研究の

3つの機能があるとされていた。一人の主任教授がその3役の全ての責任者として講座を運営していた。東京慈恵会医科大学では、1992年から大学改革が行われ、講座と診療部の分離が進められた。そして、その延長線上で、1996年に講座と教育の分離が行われたわけである。現在、診療の責任体制は診療部、教育の責任体制はコース・ユニット制となり、それぞれの責任体制が確立しつつある。東京慈恵会医科大学の教員は、研究面では講座主任教授、診療面では診療部長、教育面ではコース・ユニット責任者、そして管理・運営面では〇〇委員会委員長が上司であり、それぞれの組織で個人の力を発揮する体制ができている。余談ではあるが、かつて片岡千恵蔵が主演した「多羅尾伴内：7つの顔を持つ男」という映画があった。従来の講座制では、教員の教育、研究、診療活動の全てが一人の主任教授によって支配されていて、個人の能力を多方面にわたって引き出す体制になっていなかった。教員には多様な能力が潜在している。この多様な能力を複数の上司が異なった価値観のもとで引き出す組織こそ、組織にとっても個人にとっても働きやすい組織となる。

2. 学生評価へのあり方への疑問

講座と教育を切り離し、教育責任体制としてコース・ユニットを設定した時、学生評価をどのように行うかが問題となった。従来は各講座が100点を持ち、学生の科目合格の決定権を持っていた。各講座は独自の学生評価を行っていた。

高等教育に限らず、わが国の学生評価は、科目担当者が「私が教え、私が試験問題を作り、私が採点し、私が合否判定をする」というものである。我々はこれを「教育の密室性」と呼びたい。教育の場面では、明らかに学生・教師関係は上下関係である。これは医学部独特なのかもしれないが、他講座が行っている授業を別の講座の教員が聞きに行くことを好まない傾向がある。さらに、講座が専門分野別であるため、別の講座（非専門家）がその講座（専門家）のやっている授業に口を挟むことができなかつた。これらの傾向は教育の密室性をさらに強固なものにしていったと考えら

れる。教育学、特に教師養成では、教師に最も必要な能力は、Reflective practice (内省的実践) であるといわれている。これは、教師が密室の中で、同僚評価を受けにくい環境で、学生という立場の弱い人たちに授業を行うため、教師が自らの間違いに自らが気づかなかつた場合、間違つた行動を永遠に続ける可能性があることを戒めているのである。私の大学のクラスサイズは100名である。教員が自らの間違いに気づかず、授業を行い、そして試験という強力な武器でその間違いを学生に強要した場合、その教員は「悪の100馬力」となる。特に、医学部のように専門職業職者を育てる大学ではこのような間違いがあつてはならない。従来の教育の密室性は、教育者=評価者に起因していると我々は考えた。そこで、カリキュラム改革の時に、学生評価そのものを変えていくことにした。

大学教育の質を確保するには、教育者≠評価者の世界が必要である。これを一つの学部の中で行うには、外部試験官制度を導入するか、または教育責任組織と評価責任組織を分けるしかない。東京慈恵会医科大学では後者のシステムを採用した。それが「総合試験システム」である。

3. 東京慈恵会医科大学総合試験システム

東京慈恵会医科大学では科目群としてのコースを設定している。総合試験とはこのコースの総括的評価の一つである。コースの総括的評価はセミスターごとに行われる(年2回)。1回の総合試験の試験範囲は10数単位から16単位となる。コースの教育はコース責任者の権限であるが、これとは別にコースごとに「総合試験委員会」が組織される。コース責任者と総合試験委員長は兼任しない(教育者≠評価者)。総合試験委員長は総合試験のデザイン(どのユニットから何問試験問題を出すか、その形式(多肢選択問題、短文論述問題、長文論述問題など)をどれにするかなど)を決め、コース内の授業担当者一人ひとりに試験問題作成を依頼する。授業担当者はその依頼に沿って試験問題を作成し、総合試験委員会に提出する。総合試験委員会ではコース内の教員が作成した試験問題を全て査読し、問題の適正度、コースの学習目標との整合性、

そしてその試験問題に間違いがないかを検討する。疑問点がある場合は、作問者に問い合わせるだけでなく、必要があれば委員会で訂正する。総合試験委員会では、全委員が全問題を査読する。従つて、問題は非専門家の目も通ることとなる。このような査読を行い、選定を受けた問題だけが学生の試験に出題される。試験の実施も総合試験委員会が行う。試験監督責任者は総合試験委員があつた。試験問題の60%から70%は多肢選択問題であるので、その採点はコンピュータが行い、論述試験は1問ずつ作問者に依頼する。この時点で一切の点数調整は行われぬ。多肢選択問題、論述問題の点数は総合試験委員会で集計され、試験問題は評点データの統計を取り、事後評価を受ける。正答率の悪い問題、識別指数(全体の成績の良い学生がこの問題を正答し、成績の悪い学生がこの問題を誤答する比率)の低い問題がピックアップされ、その原因が追究される。これらの低正答率問題や不適切問題を採点に加えるかどうかは総合試験委員会で討議される。総合試験では、実施後に学生に問題と正答が公開される。学生からも、試験問題に対してクレームが文書で総合試験委員会に提出される。総合試験委員会は学生からのクレーム一つひとつに回答し、その問題を採点に加えたかどうか通知する。そして、最終的な学生評点が総合試験委員会で決定される。すなわち、学生評点の決定には教育者は一切かわらないシステムとなっている。

試験が終わり、学生の成績が確定した後、低正答率問題や不適切問題作成者には、総合試験委員会から次のような書類が送られてくる。「貴殿の試験問題は、正答率(識別指数)が低すぎました。このような結果になった原因について総合試験委員長宛にレポートをご提出ください」

我々は、客観的試験は冷徹なまでに客観的であるべきだと考える。試験問題の適性を事前にチェックし、試験後もその学生の回答パターンを分析して、そしてその結果を記録する、この総合試験システムを「厳格な成績判定」を保証し、学生が次のステップ(進級)に進むためのレディネスを判定する有効なツールであると考えている。

余談ではあるが、医学部の中には、毎週試験を行い、学生を試験で縛り付けて勉強をさせようとする人たちがいる。東京慈恵会医科大学では年2回以上の試験は許していない。かつて、寺山修司の映画のタイトルで「書を捨て、街に出よう」というのがあった。医学部は医師養成を目的にしている。試験で学生を縛りつけ、大学の言うことを聞いていれば進級できるという受け身な学生を育てるのではなく、学習する責任を学習者に自覚してもらい、学習を自己コントロールする習慣を身につけてもらい、さらに書を捨て、街に出て人間としての成長をしてもらう時間も作るべきだと思う。そのためには、カリキュラムが効率性を考えられて組み立てられている必要がある。

4. 評価を大学が管理するということ

総合試験システムでは、試験問題を教育組織とは異なる組織（総合試験委員会）が管理することとなる。評価とは、授業担当者が何を学生に教えたかを示す証拠である。試験問題を管理することで、各授業担当者が、大学が決めた教育目標に沿った授業をしたかどうかを監督することができる。授業担当者から提出された試験問題に偏りがある場合は、それを次年度に是正するよう要請する。かつて、教員によっては、自分の得意分野のみを講義し、それ以外の部分を講義しない人がいたが、このシステム導入によって、教員の独善は許されなくなった。大学は授業の1コマ1コマを管理することはできないが、試験問題を管理することで、行われている授業内容を管理することができる。専門職業職者養成に、偏った知識の伝授は危険である。大学は大学で行われている教育に責任を果たす努力が必要である。評価を大学が管理することで、Structured Curriculumを実現することが可能となる。

このシステムは教員一人ひとりにその人の教育責任を明示することにもなった。総合試験システムでは、1997年以降、学内で実施された全ての試験問題を試験問題サーバに蓄積している。何年のどのコースで誰が出題した問題なのか、そしてその問題の学生の回答パターンと模範解答が全て残されている。現在ではこの過去問題サーバを学生に公開している。学生は学内

LANで全ての問題を検索することができる。このシステム（Web-based Training：WBT）は学生が自ら自分の弱点や学習困難点を是正するために準備したものであるが、試験問題を作成者の属性とともに公開することで、試験問題作成者の教育責任をも示す結果となった。

5. 評価とは学生へのメッセージ

評価とは学生へのメッセージでもある。評価で何を求めるかで、学生の学習活動が規定される。評価で知識の量だけを求めたとすると、学生は知識の獲得のみに専念することになるだろう。評価で問題解決、技能、態度まで求める工夫をしたら、学生もその求めに応じ、求められた能力の獲得に努力する。東京慈恵会医科大学では、総合試験システムの確立によりペーパー試験での客観性が十分確保されたので、従来の方法では測れなかった能力を測る新たな評価法の確立を試みた。

問題解決能力を評価するために三段ジャンプ試験を導入した。これは、学生に事例を与え、その後2時間ほどの自己学習時間を与え、学生が自由にその課題について調べ、その後、どのように情報を集め、その事例をどのように解決しようとしたかを発表してもらうものである。プレゼンテーション試験も導入した。学生がコース内の学習内容に関連したテーマを自ら設定し、それを口頭試験教員の前でプレゼンテーションするものである。また、高学年には実技試験も取り入れた。これらの評価は、必ずしも客観的な点数とはならない。担当教員の主観が入り込む評価とならざるをえない。

評価は客観的でなければならない、という風潮がある。私はこれを「客観性伝説」と呼びたい。評価にとって、客観性が重要性より上位とは考えない。評価には、学生が伸びていく時の支援機能もある。成績判定に利用しない評価もある。例えば、ある学習プロセスの中で、学習者が自分の誤解に気づくために評価を行う場合がある（形成的評価）。特に高等教育では、学習の責任は学習者にある。学習者が自ら成長していく時に、評価という方法で学習者に気づきを与え、学習者が自ら是正していくことを促す方法は、生涯学習者

養成の手法でもある。今までの教育では、「知っていること」(Knowledge)ばかりに力点が置かれていた。専門職業職者教育ではようやく、「できること」(Competence)の教育にシフトしつつある。医学教育では、客観的臨床能力試験 (Objective Structured Clinical Examination : OSCE) と呼ばれる実技試験が全医学部で実施されている。これは、診察技法や面接ができるかどうかを試験するもので、まさに、Competenceの試験といえる。しかしながら、「知っていること」や「できること」で大学教育は足りるのであるか。

余談ではあるが、私が医学部に入学した時と、6年経って卒業する時とでは医療の世界は大きく様変わりをしていった。私は産婦人科開業医の長男として生まれ、家の家業を継ぐために医学部に入った (1975年)。この時の父の医院では年間200例を超える分娩数があり、多分、家業を継ぐ必要もあったのであろう。しかしながら、私が卒業する1981年には、父の医院の分娩数は60例程度にも少なくなり、家業としては成り立たなくなっていた。卒業時、父に言われた台詞は「継ぐ必要はない、好きなことをしなさい」であった。すなわち、急激に少子高齢化が進み、たかだか6年間で私にとっての医療環境は激変していたのである。これは医学の世界だけの話ではないと思う。社会は変化する。変化した社会が求める人材も変わる。学生が大学を出て、社会人となったとき、最も重要な能力は、「知っていること」でも、「できること」でもない、「変化に対応し、自分自身を作り変えていく力」(Capability)ではないだろうか。まさにダーウィンの世界であろう。「強いものが生き残るのではない。変化するもののみが生き残る」

Capabilityはどのように育てることができるのだろうか。その答えは、教師教育でいわれた「内省的実践 (Reflective practice)」にあると考える。内省的実践とは、自らの問題点に自らが気づき、その問題点を自ら改善すること、と定義される。この行動の第1歩は「自ら気づく」ことである。では、「自ら気づく」ためには何が必要か。それは他者であろう。他者を見

て、自分を振り返り、自分の問題点に気づき、それを改善する。Capabilityの養成には、他者を見る、すなわち、色々な価値観が存在することを学生に伝えていくことが重要である。そのためには、画一的な価値観ではなく、色々な見方、考え方があること、自分とは異なった価値観との共存の仕方を学ぶ環境が必要である。学生はまだ、十分には成熟していない。自分で気づく範囲は限定されている。そこで必要なのが「気づき」を促すサポーターである。教師には、学生に色々な気づきのきっかけを提供する役目もあるのではないだろうか。中央教育審議会答申に、「責任ある主観」による評価という言葉が載っていた。教師は「責任ある主観」によるフィードバックで学生に気づきを促し、学生の成長を支援するための評価も行うべきである。

おわりに

東京慈恵会医科大学総合試験システムでは、試験問題を通じてその教員がどのような授業をしているのかをモニターしている。授業内容に問題があると考えられた場合は、そのことを教員に通知している (フィードバック)。これは大学が行う教員評価の一つでもある。教員が作った試験問題は、模範解答だけでなく回答データも試験問題サーバに蓄積される。試験問題作成というプロの教師が行った仕事を大学が記録として保存し、試験問題の質に問題がある場合は、そのことが通知されるだけでなく、その教員の試験問題作成の能力が経時的に改善するかどうかモニターしている。

教員の最も重要な役割は、学生一人ひとりを育てることにある。学生による授業評価も、総合試験システムでのモニターも重要ではあるが、個性を持った教員が個性を持った学生にどのようなフィードバックを与え、どのように学生のロール・モデルになったか、のほうに教員の教育評価としては重要である。教員が「責任ある主観」による評価を行え、教員一人ひとりの個性が学生に見える教育環境を整えるのが大学の責任ではないだろうか。

An Attempt for Educational Assessment

—Educators are not Evaluators—

※ Osamu FUKUSHIMA

[Key Words]

evaluation, comprehensive test, structured curriculum, capability, reflective practice

Jikei University School of Medicine has employed "Comprehensive course testing system" as one of summative student assessments since 1996. In this system, educators did not involve in student evaluation (educators ≠ evaluators). This system guarantees the quality of academic education. The author mentioned about the consciousness and feature of this comprehensive course testing system in this article.

※ Office of Educational Development Jikei University School of Medicine

Computer-based Testing (CBT)

福島 統*

東京慈恵会医科大学医学教育研究室

I. はじめに

Compute-based Testing (CBT)は、臨床実習開始前の学生評価のための共用試験システムに OSCE (Objective Structured Clinical Examination: 客観的臨床能力試験) とともに用いられている¹⁾。平成17年12月からこの共用試験は正式実施となる。ここでは、共用試験になぜ、CBTが採用されたのか、そしてこれからのCBTの進むべき方向性についてまとめてみたい。

II. 共用試験でのCBTの導入の経緯

平成13年3月27日に医学・歯学教育の在り方に関する調査研究協力者会議が「21世紀における医学・歯学教育の改善方策について：学部教育の再構築のために」を出し、その別冊として「医学教育モデル・コア・カリキュラム」が発表された²⁾。この答申に、「臨床実習開始前に学生の評価を適切に行うため、その評価を各大学、各教員任せにせず、大学間で合意の上、共同で質の高い総合試験問題を作成、実施することは効率的である。(中略)このような大学間で共用できる試験の創設は、平成11年の21世紀医学・医療懇談会第四次報告及び厚生省医師国家試験改善検討委員会報告書においても検討が要請されている。(中略)知識評価にはコンピュータを用いたCBTによる多肢選択形式を用い、診察技能や態度を評価する試験では、評価方法や評価基準を統一化し、必要に応じて外部試験委員をも加えた客観的臨床能力試験(OSCE)を用いることが望ましい。」との記載がある。この答申が出される少し前の平成12年11月に「臨床実習開始前の学生評価のための共用試験システムに関する専門委員会」が作られ、共用試験をどのように実施していくかについての検討が開始された。専門委員会では当初から「紙と鉛筆」による試験ではなく、コンピュータ試験(CBT)の導入が大勢を占めた。医学分野では、医師国家試験の

ようにカラー写真、エックス線写真のような視覚教材が試験問題として多く使用されている。これを全国8千人以上の医学生に行うには膨大な印刷費が毎年かかることになる。試験実施経費のうえで、CBTは魅力的なものと考えた。また、共用試験は医師国家試験のような国が行う試験ではなく、大学が主体となって行う試験である。その理念が試験実施形態にも表現されるべきであると考えられた。従って、この試験は全国一斉に行うものではなく、各大学がそれぞれの大学のスケジュールに沿って実施できなければならない。そのためには「紙と鉛筆」による試験では不可能であるとの結論に達した。

当時、医学部で試験問題管理にサーバシステムを使っていたのが東京慈恵会医科大学であった。慈恵医大総合試験システムでは、問題作成(入力)、試験問題管理(審査と訂正)、試験問題出力(編集、印刷)、多肢選択問題の採点、試験後の解答データの解析、試験後の試験問題のプールがサーバシステムとして出来上がっていた³⁾。しかし、このシステムは試験実施が「紙と鉛筆」であり、回答はマークシートを用いるものであり、共用試験に用いるには不十分なものであった。そこで専門委員会は、試験実施をコンピュータ両面で行うためのシステム開発を始めることとなった。

共用試験CBTシステムには、各大学の教員が試験問題を作成するための入力ツール、全国から集められた試験問題を管理するセンターサーバ、試験問題を審査・改訂するためのブラッシュアップツール、受験者を管理するシステム、審査後の試験問題から問題集を作るツール、そして試験会場で試験を実施・管理し、受験者の解答を回収する試験会場システム、センターに送付された解答データを採点、統計処理するシステムなどから構成されている。この開発の中で、コンピュータを試験に用いる利点の議論が進められた。

CBTでは、受験者毎に異なる試験問題を配付することができるが、これは「紙と鉛筆」の試験でも手間をかければできる。しかしながら、順次解答という形式はコン

*Osamu FUKUSHIMA: 〒105-8461 東京都港区西新橋3-25-8.

(2005年2月14日 受理)

ビュータの特性を利用したものとして価値があると考えた。東京慈恵会医科大学では問題解決能力の評価法として、Triple Jump Exercise だけでなく、Multi-station Examination (MSE) を平成10年から試行していた¹⁾。MSEとは問題解決能力の中でも臨床推論を測ろうとするもので、順次解答型の論述試験である。第1問目は患者の初診の場面で、緊急性の評価を行わせたり、医療面接で聞くべき内容を列記させたりする問題である。受験者はこの第1問目を解答後、第2問目に進むことができる。第2問目には医療面接で得られる全情報が記載され、一部の身体診察や検査所見が提示された上で、診断確定のための検査計画を問う形になっている。従って、第1問目の答えが第2問目に記載されているため、この試験では前の問題に戻る事が禁止されている。第3問目では、検査データが示され、その検査の所見を問う問題、第4問目は診断が示され、治療の経過が書かれている。そして最後の質問はこの病態を基礎医学の知識で説明させるものである。MSEでは、実際の臨床の場面と同じ様に、受験者自らが情報を集め、集めた情報を組み上げてその患者の病態を推論していく形式となっている。この例でも分かるように、順次解答形式を用いることで知識を使う能力、すなわち問題解決能力を測ることの可能性が広がる。しかしながら、慈恵医大ではMSEを「紙と鉛筆」で行っているため、この試験の実施には強面の教員が試験監督に就くことが絶対条件となっていた。共用試験CBTシステム開発では、この順次解答型の形式の多肢選択問題をコンピュータ画面で、しかも受験者毎に異なる問題を出題できるシステムを目指した。現在、これは順次解答4連問形式(タイプQ)として完成している。順次解答4連問形式について、平成16年度医学系共用試験CBT問題作成マニュアルに次のように説明されている。「診療参加型臨床実習(クリニカル・クラークシップ)で学生がチームのメンバーとして一定の役割を果たすためには、診断や治療の決定に必要な情報を収集する能力や病態を解釈する能力を臨床実習開始前に身につけていることが重要です。そこで共用試験では臨床問題解決のために必要な学習が出来ているか否かを順次解答4連問形式の問題で評価することにしました。すなわち、診療参加型の臨床実習で学生に求められる能力である初診から診断にいたる過程での情報収集力とその情報の解釈力を評価できるように4連問を作問し、学生に順次にこの設問を解答させることにより、クリニカル・クラークシップの場において、ひとりで患者さんから情報を得てくるだけの基礎的な力がついているかを調べます。診断や治療の決

定はチームで行なえることなので、あえてはずして、第1設問は医療面接、第2設問は身体診察、第3設問は検査に限定し、診断や治療を問う設問はいれないことにしました。診断や治療を問いたいときはタイプAで出題することが出来ますので、タイプQ形式の作問ではこの原則を守ってください。また第4設問では必ず提示した症例に関連した病態生理、解剖、疫学など、学生が臨床の場で学ぶときに必要な基礎的な知識を問う設問にしてください。この設問には基礎の学習をするときにも、常に臨床との結びつきを考えるとという学習の方向性が示されています。]

タイプQはわが国独自開発の試験問題形式である。このように共用試験CBTはコンピュータ試験の特性を生かすよう開発が進められた。そして開発は今も続けられている。

Ⅲ. これからの CBT

CBTでは、今までのトライアルで使われた試験問題をその解答データとともにサーバに蓄積している。プール化された試験問題は1万題近くとなっている。平成16年12月から行われている医学系CBT最終トライアルでは解答データがついているプール問題からも出題を行っている。過去の解答データを解析すれば、受験者毎に異なる問題を出題しても、難易度を事前に調整することができる。最終トライアルでは、項目反応理論²⁾を用いて受験者ごとの難易度調整を行う試行が行われている。今後、共用試験という大規模解答データを用いた試験についての研究が期待される。

過去に出題した問題を活用する別の試みもある。東京慈恵会医科大学では、2年次後期以降卒業試験までが全て総合試験化され³⁾、平成9年度以降は試験問題サーバにその問題が全て蓄積されている(約3万題に及ぶプールがある)。この問題を利用し、学内LANでのWeb-based Training(WBT)が平成15年度から立ち上がった⁴⁾。学生は学内のコンピュータから公開用問題サーバにアクセスできる(もちろん、今年度使用問題のサーバにはアクセスできない)。学生は学年と科目を選択し、問題数を入力すれば、過去問題による問題集が出来上がり、両面上で解答をクリックすると自動採点され、科目ごとに成績が分かるシステムである。学生はいつでも、自分の能力を自己診断できる環境となっている。このシステムにより、学生は自分の弱点を知り、自分の学習を自分でモニターすることとなる。学習者が自分の学習をモニターすることは、生涯学習者として必要な能力といわれている⁵⁾。試

験をコンピュータ化することで、このようなトレーニングシステムも容易に作るができるようになる。学生はもちろん、このシステムで今までの出題問題全てを電子データとして入手することができる。学生は電子データとしてしか入手できない。すなわち、学内のコンピュータでは直接印刷できないように工夫している（紙の消費量を抑えるため）。学生がCDなどにそのデータをコピーして利用することとなる。過去問題を利用することに関しては色々な議論があると思うが、効率的に覚えていくにはこのようなシステムは時間節約になると考えている。学生が必要なことを効率的に覚える学習環境を作れば、余った時間を学生は覚えること以外の学習に振り分けることができる。すなわち、人間としての成長のための時間に。

試験問題をサーバに蓄積し、それを再利用できる環境では、試験問題作成者の責任も問われることとなる。過去問題は正答率や識別指数だけでなく、作成者の氏名も公開されている。試験は学生にとっては進級を判定される重要事項であるが、教育の立場からみれば、学習目標の明示であり、学生の学習意欲をそそるものでもある。評価が学習者の学習スタイルを規定する”。試験問題作成者の責任は大きい。従来までは試験問題が集積されることも、公開されることもなかったが、試験をコンピュータ化することでこのように試験問題を蓄積、公開することが容易となる。教員は、かつて自分が作ったプロダクト（試験問題）に責任を持つこととなる。教員は教育のプロである。プロが作ったプロダクトを人に見られることを拒むプロがどこにいるだろうか。

IV. おわりに

コンピュータを使えば、動画や音声を使った試験が可能となる。共用試験は臨床実習で学生が患者さんから学べる準備状況（レディネス）が整っているかどうかを測る目的がある。まさに、Real Patientsの前のPaper Patients (Cyber Patients) での評価が必要である。動画や

音声を試験問題に使える CBT が完成すれば、さらに学習効果の高い臨床実習への道が開かれる。

今までの CBT には、試験実施の経済性、試験問題の適正度の確保、WBT などによる学習環境の整備、タイプQなどの新形式問題の出題、そして Cyber Patients による学習と評価、試験問題作成者の責任の明示など、の利点がある。これからも新しい CBT の利点が開発されていくことと思う。しかし、CBT ではできないことがあることを忘れてはならない。学生にとって、最も重要な教材はロール・モデルである。CBT によって、いくら学習環境を整備しても、学生が医師としての誇りやプロとしての誇りを持たなければ何の意味にもならない。CBT は単なる道具でしかない。CBT は、教員の教育負担を軽減し、教員に時間を与え、その時間を学生に接する時間に当てるためのものでなければならない。教育とは「もの」がするのではない、「人」がする活動である。

文献

- 1) 福島統. 臨床実習開始前の学生評価のための共用試験システム CBT と OSCE. 医学教育 2002;33(2):83-7.
- 2) 医学・歯学教育の在り方に関する調査研究協力者会議. 21世紀における医学・歯学教育の改善方策について：学部教育の再構築のために. 2001.
- 3) 福島統. 東京慈恵会医科大学における総合試験について. 大学と学生 2001;438:41-4.
- 4) 古谷伸之, 福島統, 阿部俊昭. 探求学習による問題解決能力開発のための患者診療モデルを用いたテュートリアル教育. 医学教育 2002;33(1):21-30.
- 5) 仁田善雄, 前川真一, 柳本武美, 前田忠彦, 奈良信雄, 石田達樹, 福島統, 齋藤直彦, 福田康一郎, 高久史磨, 麻生武志. 項目反応理論を用いた共用試験医学系 CBT トライアルの解析. 医学教育 2004;35(Suppl.):70-1.
- 6) 福島統, 飛内峰和, 綱川ルリ子, 南波勉, 小松一祐. 慈恵医大総合試験システム -Web-based Training への道- [学会報告]. 平成16年度大学情報化全国大会. 2004年9月. 東京.
- 7) 米国学術研究推進会議編著(森敏昭, 秋田喜美子監訳). 21世紀の認知心理学を創る会訳). 授業を変える：認知心理学のさらなる挑戦. 京都:北大路書房;2002.