

卒前医学教育の新しい動向 3：共用試験 CBT^{*1}

福島 統^{*2}

1. モデル・コア・カリキュラムと共用試験の意義—診療参加型臨床実習を目指して

共用試験（臨床実習開始前の学生評価のための共用試験システム）では computer-based test (CBT) も objective structured clinical examination (OSCE) もモデル・コア・カリキュラムを出題基準としている。モデル・コア・カリキュラムと共用試験というこの2つの動きは、診療参加型臨床実習（クリニカル・クラークシップ）の実現を目指したものである。診療参加型臨床実習では、医学生は指導医の監督のもとに、診療チームの一員として実際の患者診療に従事する。そのため、一定の医行為が許容されるだけでなく、医学生にも一定の患者診療への責任が生じることとなる。従来の臨床実習は、臨床見学型（医学生は医師が行う診療行為を見学するのみ）、模擬診療型（実際の診療とは別に、患者さんの協力を得て医学生が模擬診療を行う）であった。しかしながら、臨床実習の教育効果として診療参加型が極めて優れていることは明らかで、基本的臨床能力養成には欠かすことのできない学習環境である。この環境が整っていないことがわが国の医学教育が国際水準に達していないといわれる理由となっている。診療参加型臨床実習は、北米、イギリスの医学教育の影響を受けているアジア諸国でも一般化している。アジア諸国に比べても、日本の医学教育は明らかな遅れをとっているのである。

診療参加型臨床実習については、すでに1987年の「医学教育改善に関する調査研究協力者会議」

最終まとめ¹⁾に、臨床実習を学生参加型にすべきであるとされ、1991年の「臨床実習検討委員会」最終報告²⁾では、指導医の監督・指導のもとで、患者さんの了解をいただいたうえで、さらに学生が臨床実習開始前までに基本的知識・技能・態度を身につけているという条件が揃えば、医師免許を持たない医学生が医療行為にかかわることの違法性が阻却されるという見解が示された。この報告にある「学生が臨床実習開始前までに基本的知識・技能・態度を身につけているという条件」を評価するために、1999年の「21世紀医学・医療懇談会第4次報告」³⁾に臨床実習への「適切な進級認定システム」の提案があり、それを受けて2001年の「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—」⁴⁾に医学教育モデル・コア・カリキュラムの提示と「臨床実習開始前の学生評価のための共用試験システム」の提案がされた。

診療参加型臨床実習を行うには、診療現場の安全性を確保するために、医学生に一定の知識、技能、態度が身につけている必要がある。それだけでなく、学生は臨床実習の場で、「患者さんから学ぶ（患者こそ最高の師）」ことが求められている。この患者さんから学ぶ力がなければ、学生は、臨床実習という膨大な時間を無為に過ごすこととなる。共用試験とモデル・コア・カリキュラムは、学生が臨床実習開始前までに着実に臨床実習で患者さんから学ぶ能力を身につけ、患者さんの協力のもとで学生が医師としての素養を身につけていくためのものである。

2. コンピュータ試験開発の理由

2002年の「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—」⁴⁾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

^{*1} The Common Achievement Test System: Computer-based Test (CBT)

キーワード：CBT, 教育における質の保証, 学習評価

^{*2} Osamu FUKUSHIMA 東京慈恵会医科大学教育センター

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)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2001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、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いく試験である。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。したがって、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。東京慈恵会医科大学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でき上がっていた⁵⁾。しかし、このシステムは試験実施が「紙と鉛筆」であり、解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画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には、受験者ごとに異なる試験問題を配信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ピ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、multi-station examination (MSE)を1998年から試行していた⁶⁾。MSEとは臨床推論能力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。したがって、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ように、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、すなわち臨床推論能力を測ることが可能となる。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、しかも受験者ごと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、2004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がされている。「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できていないかを順次解答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な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、あえてはせずして、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設問では必ず呈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。

3. CBT 問題の流れ

共用試験 CBT では、第1回トライアルから試験問題作成を全国医学部に依頼し、提出された試験問題を共用試験実施機構内の委員会でも専門分野の異なる委員の目でブラッシュアップした問題のみを受験者に試験している（ブラッシュアップ後の試験問題の採択率を表1に示す）。出題され、成績データを付加された問題はさらに、事後評価委員会での試験問題としての適正について検証が

行われ、プール問題として蓄積される（プール問題累計数を表1に示す）。プール問題は常時見直される体制が整いつつある。このように共用試験 CBT では質の高い試験問題が供給されている。

4. CBT トライアルの経過

CBT だけでなく、共用試験の今までのトライアルの経過については報告書⁷⁻¹⁰⁾に詳しく記載されている（これらの報告書は資料としての価値が高い）。

共用試験の正式実施（かつては「本格運用」と呼ばれた）は2005年12月からである。それまでのトライアルについて表1と表2にまとめた。第1回 CBT トライアルは、2002年1月から5月に77医学部で実施され、出題数は単純5肢択一形式（タイプA）100題（2ブロック）であった。第2回目以降は全国80医学部すべてが参加している。

試験形式の変遷を表2にまとめてある。第2回トライアルでは順次解答2連問形式（タイプW）と順次解答4連問形式（タイプQ）が導入された。タイプWとは、1つのテーマについて臨床と基礎の問題をセットにしたものである。その例題は2003年の報告書⁷⁾に載っている。例題では熱傷でのショックをテーマとし、第1設問で熱傷の程度の診断を、第2設問では同じ症例での病態生理についての基礎知識を問う形式となっている。基礎・臨床の統合を試験問題として表現しようとする意図がある。順次解答4連問形式（タイプQ）は前述したように、臨床推論能力を測ろうと作られた問題形式である（関節リウマチをテーマにした例題が報告書⁷⁾に示されている）。第3回トライアルでは extended matching items と呼ばれる北米、英国で多用されている多選択肢問題の導入が行われた（分娩時の異常出血をテーマにした例題が報告書¹⁰⁾にある）。多選択肢型は2連問形式と4連問形式の2種類が出題された。第3回トライアル後、出題形式についての議論があり、最終トライアルでは、ブロック5では多選択肢2連問形式のみを、そしてブロック6ではタイプQのみを出題することとなった。この議論では、タイプWを順次解答の2連

表1 CBTの結果

トライアル回数	実施時期	平均正解率	正答率標準偏差	平均難易度	難易度標準偏差	公募問題数	採択問題数	選択率	プール問題累計数
第1回トライアル	2002年1月～5月	55.90%	10.2	55.7	2.1	9,919題	2,791題	28.14%	2,305題
第2回トライアル	2003年2月～5月	55.90%	9.4	55.9	1	9,322題	3,723題	40.00%	5,208題
第3回トライアル	2003年12月～2004年3月 2004年6月～9月	57.30%	9.8	57.2	1	7,108題	3,803題	53.50%	9,083題
最終トライアル	2004年12月～2005年3月 2005年6月～9月	59.80%	10.1	59.5	0.9	7,219題	2,440題	33.80%	10,834題以上

注1. 最終トライアルでの採択率は順次解答型（21%）や多選択肢型（34.1%）を含む。

注2. 最終トライアルでのプール問題累計数は前期修了分のみの加算のため、以上と記した。

表2 CBTの変遷

	出題問題数	試験時間	出題内容	変更点
第1回トライアル	100題	2時間	単純5肢択一形式（タイプA）のみコアカリ出題割合 A：5題，B：20題，C：40題 D：10題，E：15題，F：10題	
第2回トライアル	300題	6時間	単純5肢択一形式：240題 出題比率 A：約5%，B：約20%，C：約40% D：約10%，E：約15%，F：約10% 事例提示順次解答2連問形式：40題 事例提示順次解答4連問形式：20題	受験者管理の導入 新形式問題の導入（順次解答形式） ブロック5を順次解答2連問形式（タイプW） ブロック6で順次解答4連問形式（タイプQ）を出題 追再試験が可能 CBT事前体験ツールの配布 学生向けパンフレットの配布 CBTモニター派遣 MDセンターの実施サポート 試験開始許可キーの導入 CBT実施ログの回収
第3回トライアル	320題	6時間	単純5肢択一形式：240題 出題比率 A：約5%，B：約20%，C：約40% D：約10%，E：約15%，F：約10% 順次解答2連問形式（タイプW）：20題 多選択肢2連問形式（タイプL）：20題 順次解答4連問形式（タイプQ）：20題 多選択肢4連問形式（タイプR）：20題	年間2回実施（12月～3月と6月～9月） 多選択肢形式の導入（ブロック5と6） ブロック6を20題から40題と出題数を増加 順次解答2連問形式の作成を各大学に依頼 順次解答4連問形式の作成を各大学に依頼
最終トライアル	320題	6時間	単純5肢択一形式：240題 出題比率 A：約5%，B：約20%，C：約40% D：約10%，E：約15%，F：約10% 多選択肢2連問形式（タイプL）：40題 順次解答4連問形式（タイプQ）：40題	ブロック5は多選択肢2連問形式のみ出題 ブロック6は順次解答4連問形式のみ出題 タイプWとタイプRは出題しない 多選択肢4連問形式の作成を各大学に依頼

注1. 東京医科歯科大学医歯学教育システム研究センター（MDセンター）が2002年4月に全国共同利用施設として設置された。

注2. 正式実施は2005年12月から開始された。

問としてセットで出題する意義について議論となり、タイプAの問題の質の向上をはかり、タイプAで基礎・臨床統合問題を作成、出題する、タイプQの出題数を増やすことで受験者に臨床推論能力開発の重要性を伝える、北米・英国で一般化しつつある多選択肢問題を増やすことが結論された。この考え方は正式実施に受け継がれ、正式実施では最終トライアルと同じで、ブロック5で多選択肢2連問形式40題、ブロック6で順次解答4連問形式40題が出題されている。

最終トライアルでは、プール問題と試行問題の両方が出題されている。プール問題は過去に出題され、受験者の解答データが付加されている。この解答データは項目反応理論をもとに、その問題の難易度が計算されている。この手法を用いることで、各受験者に出題される問題セットの難易度がほぼ一定となっている（最終トライアルでは問題セットの平均難易度が59.5で、その標準偏差は0.9である）。最終トライアルから評点はプール問題のみの採点となっている。プール化された試験問題は1万題を超えた。

5. CBTの今後

CBTの今後の課題としては、現行システムの安定化と改修、コンピュータ利用の利点の応用（動画、音声などの試験素材への利用など）の検討、出題形式の検証、項目反応理論を用いた試験問題の研究、継続的プール問題管理などが挙げられる。

評価とは学生へのメッセージである。共用試験は臨床実習開始前の学生にどれだけの能力を求めているのかの指標である。CBTは今後、この問題と直面していく必要がある。すなわち、学生に

「臨床実習開始前」にどのような能力をどれくらいの深さで求めるのか、を明確にしていかなければならない。さらに、国民に対しては、これだけの能力を保証しますので、学生が臨床実習に行ったときには、学生の教育にご参加ください、といえるものでなければならない。各大学の英知を集め、この難題に向かっていく必要がある。

文 献

- 1) 医学教育の改善に関する調査協力者会議。最終まとめ。1987年9月7日。
- 2) 厚生省：臨床実習検討委員会最終報告。1991年5月13日。
- 3) 21世紀医学・医療懇談会第4次報告。21世紀に向けた医師・歯科医師の育成体制の在り方について。1999年2月26日。
- 4) 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年3月27日。
- 5) 福島 統。慈恵医大における総合試験について。大学と学生2001; 438: 41-4。
- 6) 古谷伸之, 福島 統, 阿部俊昭。探索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: 21-30。
- 7) 臨床実習開始前の医歯学生の臨床能力評価方法に関する調査研究。平成13年度～14年度科学研究費補助金（特別研究促進費）成果報告書（課題番号：13800002）。
- 8) 臨床実習開始前の共用試験—平成15年。共用試験実施機構。全国医学部長病院長会議。2003年8月。
- 9) 臨床実習開始前の「共用試験」第2版（平成16年）。共用試験実施機構。医歯学教育システム研究センター。2004年11月。
- 10) 臨床実習開始前の「共用試験」第3版（平成17年）。社団法人医療系大学間共用試験実施評価機構。医歯学教育システム研究センター。2005年9月。