

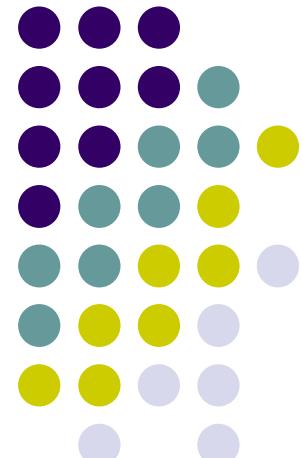
医療系大学間共用試験について



公益社団法人医療系大学間共用試験実施評価機構
CATO : Common Achievement Tests Organization

理事・事業部長

石田達樹

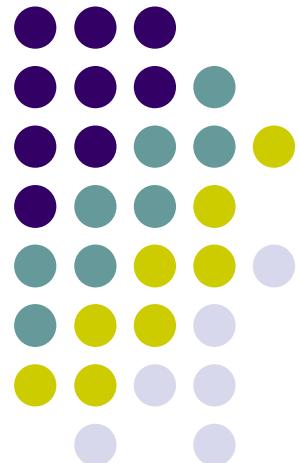




本日のおはなし

1. 共用試験導入の経緯
2. 共用試験の現況
3. CBTの概要
4. CBT導入の概要等
5. CBTトライアル
6. IRTの適用

共用試験導入の経緯





医療系大学間共用試験実施評価機構

○目的 (会員は学部長・学長 公式HP <http://www.cato.umin.jp/>)

臨床実習において患者の診療に参加する学生の知識、態度及び技能が標準的な水準に到達していることを評価する共通の評価試験（以下「共用試験」という。）の実施に関する事業を行うとともに、医療系大学間における教育の質の向上と充実を図るための事業・学術研究・啓発・普及活動を行う。もって、優れた医療人を育成し、国際的に高く評価されている我が国の医療の向上と国民の福祉の増進に寄与することを目的とする。

○事業

- 1 医療系大学等の教育における学生の学習到達度を判定するための共用試験の実施及び評価に関すること
- 2 共用試験の問題及び課題の作成に関すること
- 3 共用試験の問題、課題及び成績等の管理に関すること
- 4 共用試験の内容及び方法の改善を図るための組織的な研究、研修に関すること
- 5 医療系大学等の教育内容とその評価の発展充実を図るための学術研究活動、研修事業、資料収集及び啓発・普及活動等の実施に関すること
- 6 その他機関の目的を達成するために必要な事業

○沿革

2001年2月 共用試験トライアル開始

2003年4月 任意団体 共用試験実施機構を設立

2005年3月 文部科学省より社団法人医療系大学間共用試験実施評価機構設立認可。

医学部・医科大学80大学、歯学部・歯科大学28大学。(2007年度奥羽大学歯学部加入、2016年度東北医科薬科大学加入、2017年国際医療福祉大学加入により**医学部・医科大学82大学、歯学部・歯科大学29大学**)

2005年12月 第1回共用試験(CBT & OSCE) 正式実施開始

2014年4月 内閣府より公益認定を受け、公益社団法人に移行。

2019年2月 理事長に栗原敏（学校法人慈恵大学理事長）就任。



臨床実習開始前の共用試験導入の経緯等

医学教育改革がきっかけ

1983年：AAMC GPEPレポート (Acad. Med. 1984)

●医学教育の改善に関する調査研究協力者会議（文部省、1987年）：

- ・教育目標の明確化、カリキュラム改善、臨床実習充実と評価、卒前と卒後研修の関連等多くの提言

●臨床実習検討委員会最終報告（厚生省、1991年）：

- ・医師法との関係、医学生に許容される医行為のレベルと事前評価

1983年：GMC Tomorrow's Doctors 、 AAMC ACME-TRI

●「21世紀医学・医療懇談会報告」第1次～第4次報告（文部省、1996年～1999年）：

知識の伝授のみで、態度・技術の習得が十分行われていないという問題意識

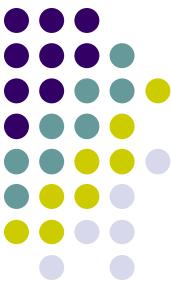
- ・適性に留意した人材選考、臨床実習の充実、教育内容の精選と多様化、適切な進級認定システムの構築 1999年：GMC Implementing Tomorrow's Doctors

●「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－（文部科学省、2001年3月）：

重要

- ・学部教育内容の精選＝「モデル・コア・カリキュラム：教育内容ガイドライン作成
- ・臨床実習開始前の適切な評価システム構築＝共用試験システムの開発

2002年：GMC Tomorrow's Doctors



共通の評価システムとしての共用試験

「臨床実習に臨む学生の能力・適正について、全国的に一定の水準を確保するとともに、学生の学習意欲を喚起する観点から、

共通の評価システムを作る

ことを検討すべきである。」

1999年2月「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」
(21世紀医学・医療懇談会第4次報告)

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
医学・歯学教育の在り方に関する
調査研究協力者会議 報告



「各医科大学（医学部）の関係者へ」

- 本会議で提案させていただいた、・・・（中略）・・・などについては、来年度中（平成13年度）には早急に学内において改善のための検討を進め、再来年度（平成14年度）のカリキュラムには着実に反映していただくことを・・・。
- 社会経済が大きく変革を遂げていく中で、その速度に応じた医学教育の改革を進めていかなければ医科大学にとって未来はやってこない・・・。



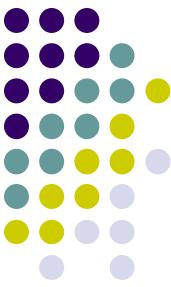
「医学生へ」

- 教育内容や方法が改善されても、受ける側の医学生の意識改革が進まなければ、成果はありえない。
- **自宅での予習を含めた自学・自習による幅広い勉学の成果に対する期待がある。・・・・**
学外での貴重な自学・自習時間が、アルバイトなどに浪費されるような状態では、この改革の成功はありえない。



「国民の皆様へ」

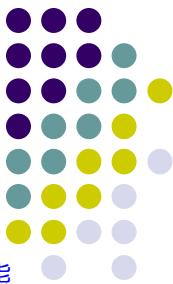
- 良い医師・歯科医師を養成することは、社会的使命であり国民だれもが望むところである。良い医師・歯科医師を養成するためには、どのように素晴らしいカリキュラムや教育体制の整備がそろっていてもそれだけでは不十分で、**医学生が数多くの臨床経験を積むことが必要不可欠である。**
- すなわち、良い医師・歯科医師の要請は、臨床実習を通じて患者中心の医療を医学生・歯学生が体得してはじめて実のあるものとなる。



「関係行政機関へ」

- 特に、今回提案したコア・カリキュラム、臨床実習開始前の評価システムについては、国家試験や卒後臨床研修との調整を図っていくことが求められる。

臨床医学教育では「臨床実習」がポイント



「人から学ぶ」：教科書に書いてあることは一般論であるが、一人の人がその病を得たとき、症状も異なるばかりでなく、その病により精神状態も、心理社会的反応も異なる。→場面に応じた思考が求められる。

臨床実習：

ポリクリニック・・・・・・・ 従来の見学型臨床実習
ベッドサイド・ティーチング・・・・ 教員視点
ベッドサイド・ラーニング・・・・ 学習者視点
クリニカル・クラークシップ・・・・

医学生が診療チームの一員となり、

指導医の監督のもと、

「定められた範囲で」 実際の診療に携わる

診療参加型臨床実習

医行為

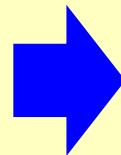
メディカル・インタビュー、身体診察、検査計画立案、
定められた範囲の検査を実施、検査結果判読、病態解析、
治療計画立案、定められた範囲で治療手技を実施など

一定の医行為が許容されるだけでなく、医学生にも患者診療への責任が生じることとなる。



違法性阻却事由の1つとしての共用試験システム

- 医師・歯科医師として資格のない学生が患者に接して医行為を行いうる不可欠な要件（**医師法第17条・歯科医師法第17条の違法性の阻却要件**：①患者への侵襲性や患者の羞恥心を十分配慮した一定の行為に限定する、②指導医のきめ細やかな指導・監督下に行う、③患者などの同意を得て行う、④事前に学生の能力と適性の評価を行う）の一つとして、事前に学生の能力と適性を評価し、質を保証することが必要。



- 臨床実習開始前に学生の能力と適性を適切に評価する
- 共同で質の高い総合試験問題を作成・実施
- 臨床実習に参加している学生が一定の能力を有することを社会に保証する共通の評価基準を設定

参考：臨床実習検討委員会最終報告（H3.5.13前川レポート：厚生省健康政策局長通知 健政発第306号、H3.6.7 文部省 医学教育課長通知 3国高医第17号）



共用試験：学生の能力と適性について全国的に一定水準を確保するための標準評価試験

医学教育モデル・コア・カリキュラム(平成28年度改訂版) 概要

- 学生が卒業時までに身に付けておくべき、必須の実践的診療能力(知識・技能・態度)を、「ねらい」と「学修目標」として明確化
- 学生の学修時間数の3分の2程度を目安としたもの
- 「医師として求められる基本的な資質と能力」として、ミニマム・エッセンスである項目を記載

2001年初版、2007年第2版、2010年
第3版、2016年第4版 発行

多様なニーズに対応できる医師の養成

A 医師として求められる 基本的な資質・能力

プロフェッショナリズム

医学知識と問題対応能力

診療技能と患者ケア

コミュニケーション能力

チーム医療の実践

医療の質と安全の管理

社会における医療の実践

科学的探究

生涯にわたって共に学ぶ姿勢

B 社会と医学・医療

集団に対する医療

法医学と関連法規

医療に関連のある社会科学領域

C 医学一般

生命現象の科学

個体の構成と機能

個体の反応

病因と病態

人の行動と心理

D 人体各器官の正常構造と
機能、病態、診断、治療

E 全身に及ぶ生理的变化、病態、診断治療

F 診療の基本

早期臨床体験実習

症候・病態からのアプローチ

基本的診療知識

基本的診療技能

C 診療参加型臨床実習開始前の「共用試験」
B T (知識) · O S C E (技能・態度)

医学研究と倫理

G 臨床実習

医師国家試験
(知識)

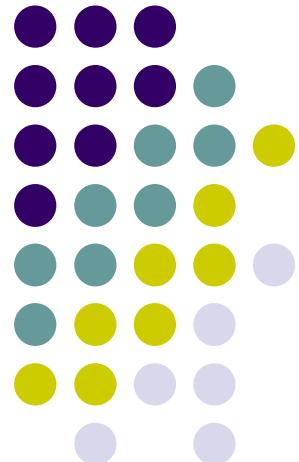
(技能・態度)

各大学の特色ある独自のカリキュラム

- 各大学が教育理念に基づいて実施する独自の教育内容(教養教育や、学生が自主的に選択できるプログラムを含む)
- 学生の学修時間数の3分の1程度

MEXT 連絡調整委員会、専門研究委員会策定

共用試験の現況



大学入学から生涯教育(制度・試験) 医学の場合 (Ver8.5)

地域医療・医師養成における都道府県の権限強化

2020(R02)年までに各種仕掛け構築

医療従事者の需給に関する検討会（中間まとめ）

- ・'16年度入学定員 9,262人
- ・'24年医師需給均衡（中位推計）
- ・'54年3.4万人の医師供給過剰
- ・'17年度で終了する医学部定員増の暫定措置は当面延長
- ・'20年以降の医師養成数は今後検討各都道府県において
- ・'19年中 医師確保計画策定、翌年実施

MEXT 連絡調整委員会、専門研究委員会

モデル・コア・カリキュラムの定期改訂（H28年度第3次改定）、英文化

厚生労働省 医道審議会医師分科会で検討

Student Doctor
統一合否水準設定

Curriculum policy

各段階での厳格な質保証の実施

Diploma policy

MHLW医師分科会
臨床研修部会

新たな専門医養成制度
日本専門医機構(JMB)発足'160704
新体制発足（吉村理事長）
…見直し開始
'170615 専門医制度新整備指針(第2版)
'180720新新体制発足（寺本理事長）

1 2

3 4

5

6

医学教育モデル・コア・カリキュラムの実践

準備教育 +各大学が教育理念に基づいて策定する
特色ある独自カリキュラム

臨床実習前医学教育

高校教育

基 本 事 項 教 育

生涯
教育

Admission policy

センター試験、入学試験
H32～大学入学希望者
学力評価テスト、入学試験
(高大接続改革)

公的化

CATO,AJMC

臨床実習前の共用試験
(CBT&OSCE)…
Student Doctor証の
発行 '16年度から正式実施

CATO

臨床実習後
(Post-CC) OSCE
'20年度から正式実施予定

卒前・卒後統合EPOC

-EPOC2-

MHLW医師分科会
医師専門研修部会

医師国家試験
'18年度から出題数減
(400問) 期間減 (2日)

修了認定、
修了証交付

医籍登録、保険医登録

医籍登録 (2回目)

MHLW 医師分科会医師国家試験改善検討部会・出題基準改定部会

一気通貫（医師養成の一貫性担保）

地域医療構想・働き方改革・医師偏在対策

分野別評価

国際基準に準拠した基準に準拠した認証評価の実施

（認証組織JACME設立）

- ・医学教育の「2023年問題」への対応
- ・医学教育の質向上
- ・世界医学教育連盟（WFME）による
日本医学教育評価機構（JACME）
の認証完了'17/4から10年間

学習評価項目によるOSCE出題範囲の調整

定期的な摺り合
わせが必須

モデル・コア・カリキュラムによるCBT出題範囲の調整
ブループリント（CATO作成）による出題割合、難易度の調整

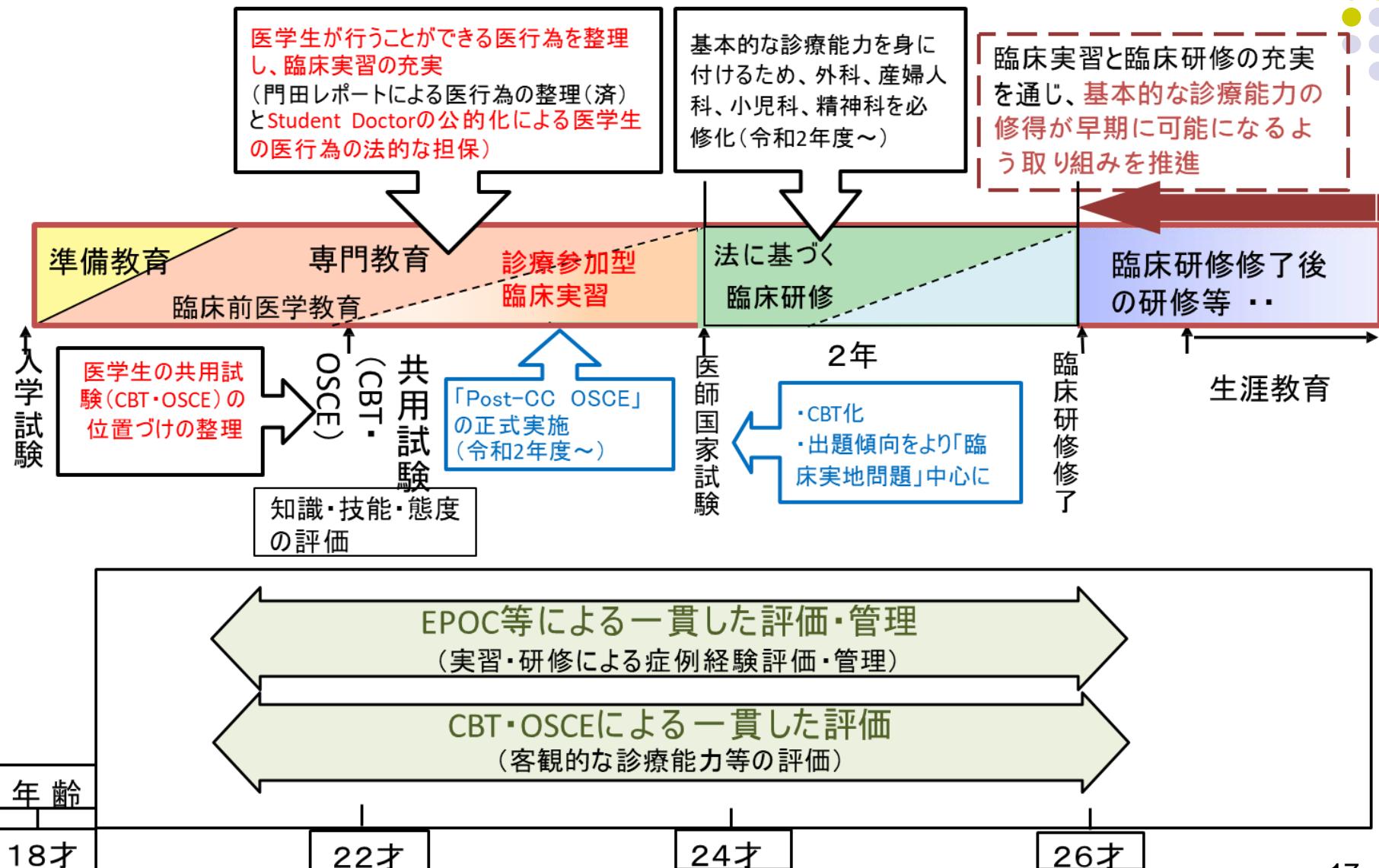
臨床研修到達目標に準拠した管理

生涯教育カリキュラム2016（日本医師会）

医師国家試験出第基準による出題範囲の調整
ブループリント（医師国家試験設計表）による
出題割合の調整

シームレスな医師養成に向けた改革全体系

厚生労働省医道審議会 医師分科会
令和元年6月19日 資料1



2 共用試験の公的化といわゆるStudent Doctorの法的位置づけについて

医師法の改正へ

(1) 共用試験CBTの公的化

- ・ **共用試験CBT(Computer-Based Testing)**は、平成17年より正式実施されており、現在では全大学で実施され、項目反応理論などの問題の精度管理の手法や評価手法が確立している。
- ・ 共用試験CBTは、日本の医学教育の中でその位置付けは既に確立されており、医師国家試験の受験要件とする等による**公的化に相当する試験**である。

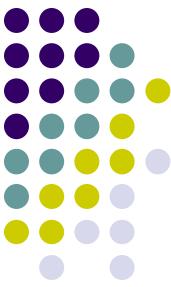
(2) 共用試験臨床実習前OSCEの公的化

- ・ **共用試験臨床実習前OSCE**(Objective Structured Clinical Examination)は、共用試験CBTと同時期の平成17年から正式実施され、その後は全ての医学生が受験するなど、**現状の医学教育の中で臨床実習前に技能と態度を試験する機会として確立**されている。
- ・ 近年、医師の倫理観や医療の安全性について、社会からより高い水準を求められることに鑑みても、**臨床実習前に一定水準の技能・態度のレベルに達していることを試験することは、極めて重要である**と考えられことから、共用試験臨床実習前OSCEは、**共用試験CBTとともに公的化すべきである**と結論づけられる。
- ・ **模擬患者**(SP: Standardized Patient)が重要な役割を果たしているが、公的化にあたっては、**全国的に取り組む組織の創設やSPに対する研修体制の整備などの検討が必要である**。
- ・ 公的化するに当たっては、公的な場においてその判断基準を協議し、より公平に判断される体制の構築について今後検討する必要がある。

(3) いわゆるStudent Doctorの法的位置づけ

- ・ **臨床実習開始前の共用試験を公的化することで、共用試験後に臨床実習を行う医学生は一定の水準が公的に担保されることから、実習において医行為を行う、いわゆるStudent Doctorを法的に位置づけることが可能となる。**
- ・ 医学生が医行為を行う臨床現場で、患者の前で各医行為が法的に認められているものかどうかを判断することが、診療参加型臨床実習の障壁になることが予想され、臨床現場で行われる診療は日進月歩に進歩していくことに鑑みると、いわゆるStudent Doctorが法に基づき行える医行為を個別に列挙することは適当ではない。必ず、医師の指導や監督のもと行われることから、**実施する行為については、指導する医師が適宜、医学生の能力と患者の状態等を勘案して判断すべきものである。**

医道審議会医師分科会報告書「シームレスな医師養成に向けた共用試験の公的化といわゆる
Student Doctorの法的位置づけについて https://www.mhlw.go.jp/stf/shingi2/0000200876_00003.html

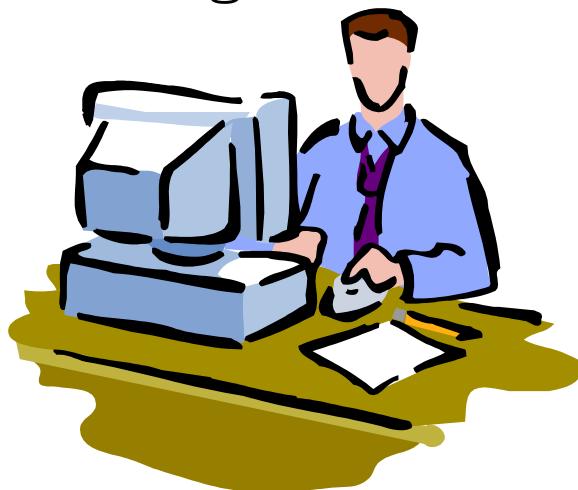


共用試験の両輪

診療参加型臨床実習開始前に修得すべき

□ 知識・（技能）

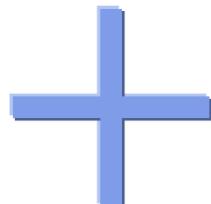
多肢選択筆記試験→**CBT**：
Computer Based
Testing



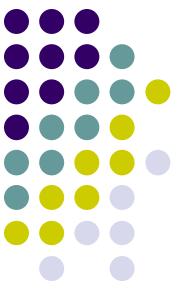
MCQ問題（320設問、6時間）

□ 技能・態度

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**を加える→**OSCE**：
Objective Structured
Clinical Examination
(客観的臨床能力試験)



医学・歯学教育のあり方に関する調査研究協力者会議最終報告(2001年3月27日)



OSCE (Objective Structured Clinical Examination: 客観的臨床能力試験)の実際

● 基本6ステーション

医療面接	10分
頭頸部	5分
胸部・全身状態とバイタルサイン	5分 (※)
腹部	5分
神経	5分
基本手技・救急	5分 (※)

● 追加ステーション

各5分
実施大学の希望に応じて、四肢と脊柱、※印の課題など最大3ステーションを追加実施可能

● 受験者の移動方法

回転式（ローテーション方式、ショットガン方式）
順次式（トコロテン方式、通り抜け方式、ワンウェイ方式）

● 認定外部評価者養成数（ST毎に認定）

平成30年度までの認定者数は延べ12,700名
毎年3回程度評価者講習会を実施
(約1,000人養成)

● 各ステーションで実施する課題毎に

課題シート、評価表、評価マニュアル、運用メモ、運用参考図、SP用シナリオ等の資料を準備

併せて 学習用DVD、評価者用DVDを作成・配布

● 模擬患者の養成

● 評価方法

チェックリスト形式の評価項目による細目評価と細目評価で評価困難な全体の流れや円滑さ、医師としてふさわしくない行為などを評価する概略評価の2本立。

● 概略評価の区分

- | | |
|---|--------------------|
| 6 | 優れている（医師と違ひがないレベル） |
| 5 | 良い（学生としては良くできるレベル） |
| 4 | 合格レベル（最低要求レベルよりは上） |
| 3 | 合否境界領域 |
| 2 | 不合格だが改善可能 |
| 1 | 明らかに不合格 |

内部評価者
(学内)

内部評価者
(学内)

機構派遣認定
外部評価者



標準模擬患者 (SP)

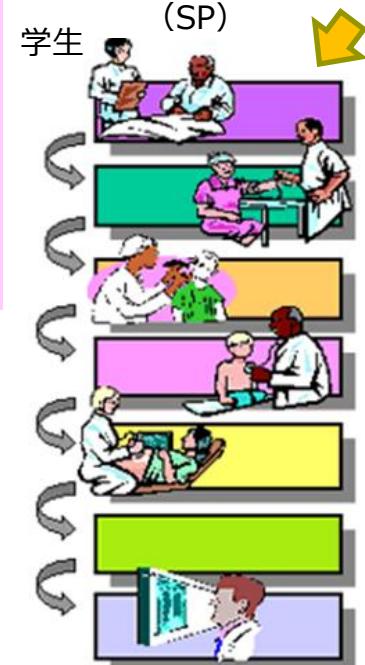
学生

【医療面接ステーションの例】

標準模擬患者
(SP)

学生

機構派遣監督者、内部・外部評価者



身体診察など

学生は各ステーションを順番に回り、基本的診療能力の評価を受け、全てのステーションに合格しなければならない。

OSCE実施風景



評価者

模擬患者はボランティア



医療面接

評価者



胸部診察

(シミュレーター使用)

学生
標準模擬患者

患者役は学生（3年次）



頭頸部診察面

学生

腹部診察

患者役は学生（3年次）

学生

評価者



外科手技・救急の例
(シミュレーター使用)



OSCEの評価者（自大学教員・他大学教員）は、評価マニュアルに則って評価表をつける

受験者は数個～十数個の小部屋(station)を順番に回り、課題(医療面接・身体診察・処置の技能など)の performanceについて評価される

標準模擬患者
(SP)



受験者

評価者



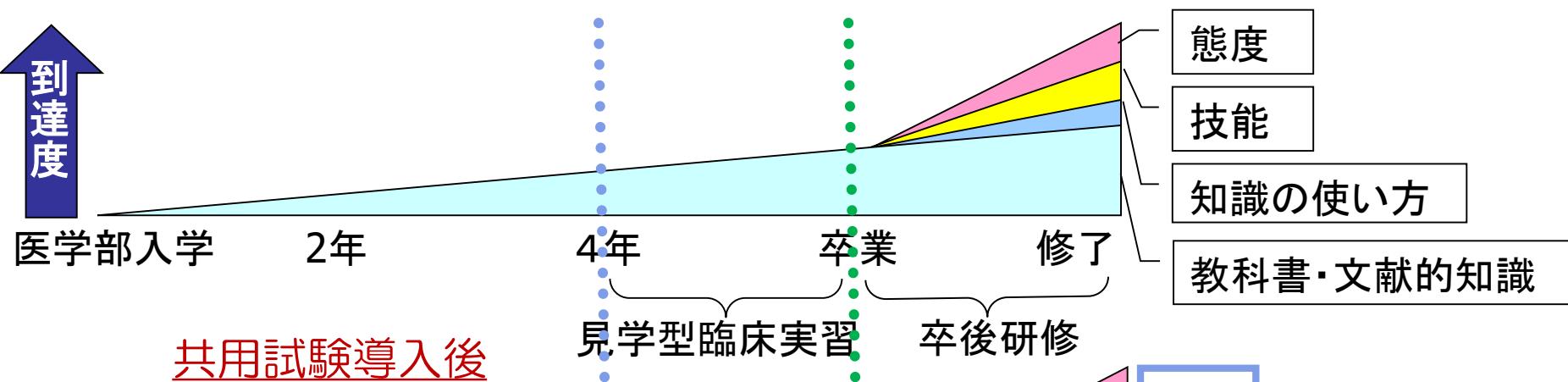
医療面接

腹部診察

診療参加型臨床実習の導入と 共用試験



共用試験導入前



臨床実習前共用試験CBT&OSCE

診療参加型臨床実習

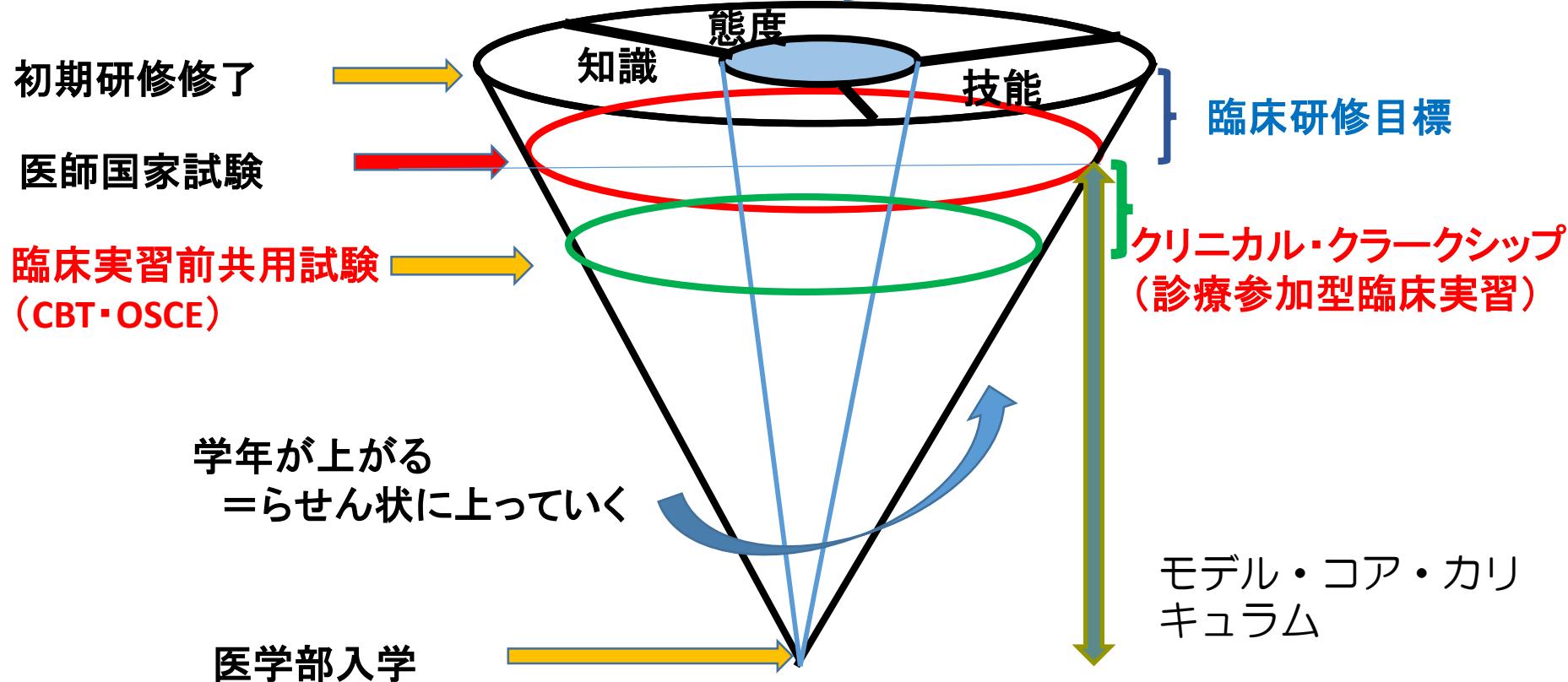
医師国家試験

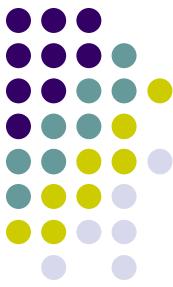
臨床実習後OSCE
Post-CC OSCE



医学部入学から初期臨床研修終了まで

研修修了時のoutcomeをEntrustable Professional Activity
(ひととおりの診療ができるレベル)とすると





共用試験の特徴

- 国家試験を除いて全医学生・歯学生に対する全国共通の標準試験の実施は日本では初めて。
- 医学系、歯学系全大学が会員となる公益法人を設置し、試験問題の作成から試験の実施、改善までを行う。
- 試験日が大学によって異なる。
- 試験問題も受験生ごとに異なる。
- 医師法、歯科医師法とも関連。
- 進級要件、合格しないと臨床実習へ進めない。
- 到達すべき学修目標は「モデル・コア・カリキュラム」に示されている。
- 実施の準備と実施当日の運営は参加大学教職員が責任者となって行う。
- CBT, OSCEの公平性・客観性・透明性を確保するために、他大学の教員が公益社団法人医療系大学間共用試験実施評価機構から機構派遣監督者として派遣される。
- 各大学教職員の多大な協力のもと運営されている。

診療参加型臨床実習開始前の到達レベル評価の仕組



全国の大学が共通で利用できる標準評価システム(共用試験)を用いて実施

- 知識の統合的理解・問題解決能力、臨床推論等の評価については、異なる時期・場所でも公平に実施・評価できる方式

→ コンピュータを用いた客観試験 CBT を選択

○ プール問題によるランダム出題方式＝システム開発、適正な試験問題の蓄積、学生の成績と試験問題のセキュリティ確保

○ 項目反応理論(IRT)の応用：問題特性の解析、出題、成績分析(能力値推測)

○ 実施本部体制構築と機構試験監督の派遣

- 態度・技能については、客観的臨床能力試験OSCEで評価

○ 技能・態度に関する学習評価項目と実施課題作成

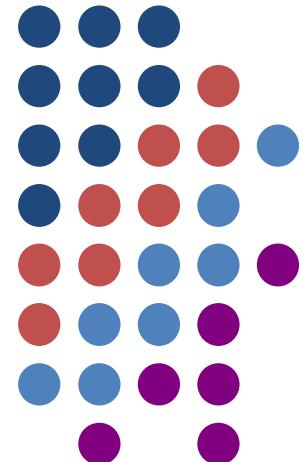
○ 評価法・評価マニュアル、教育用DVD・評価用DVD作成、評価者講習会、データ等管理システム開発、模擬患者養成、全大学での実施準備

○ 実施本部体制構築と機構試験監督・ST別外部評価者(6名)の派遣

試験の信頼性・公平性・客観性・透明性等を担保するために我が国の代表的な試験評価の専門家が参画する試験信頼性向上専門部会(※)を設置し、恒常的に活動 (※2020年4月から試験信頼性向上委員会²⁶)



CBT (Computer-Based Testing)



共用試験実施の概要（2005年12月から正式実施）



CBT (Computer Based Testing) の実際



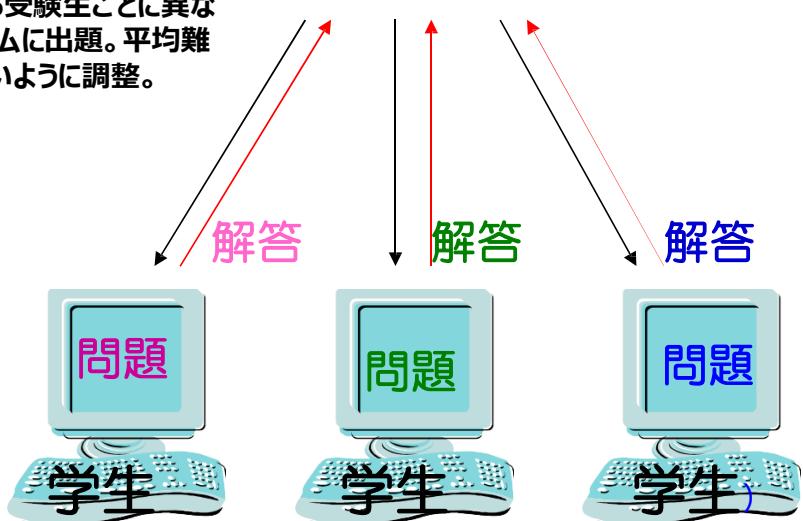
各大学
サイトサーバー

項目反応理論 (IRT:Item Response Theory) の利用

- 出題問題の事後評価 (項目特性曲線)
- 受験生毎に異なる出題セット間の難易度調整 (難易度の等質化)
- 受験生の能力評価 (IRT標準スコア)



問題プールから受験生ごとに異なる問題をランダムに出題。平均難易度に差がないように調整。



試験端末(受験者毎に問題が違う)

出題数 = 320設問/学生/1日

全6ブロック : 1ブロック60分

ブロック 1 ~ 4 : 単純5肢択一形式60設問 1問 1分

ブロック 5 : 多選択肢択一形式40設問 (鑑別診断、病態等) 1問1分半

ブロック 6 : 順次解答4連問5肢択一形式40設問 (臨床推論) 1問1分半

公益社団法人
医療系大学間共用試験実施評価機構
センターサーバー

問題

解答



Common Achievement
Tests Organization

自動採点

参加大学へ成績送付

項目特性値を
持った大量の
問題をプール

2018時点
27,000題

- 新規作成問題 (採点対象外) + プール問題 (採点対象) の組合せによる出題問題セット作成と調整作業
- CBT実施セット準備・配布
- 新規出題問題の事後評価作業
問題の特性評価、プール問題候補決定
プール問題の特性変化検討
- 共用試験実施評価機構委員によるブラッシュアップ集中作業
新規出題 (試行) 問題決定
- 各大学へ問題作成依頼
各大学でのブラッシュアップ後に新規作成問題提出

試験実施：機構派遣監督者



- ・試験の信頼性の観点から各大学に他大学の教員・機構の委員を機構派遣監督者として派遣
- ・試験の実施状況を把握し、試験が問題なく実施されたかどうかを判断
- ・試験の状況を講評し、機構派遣監督者報告書にて改善を要する事項等を報告。次年度の監督者は改善状況を確認

機構派遣監督者





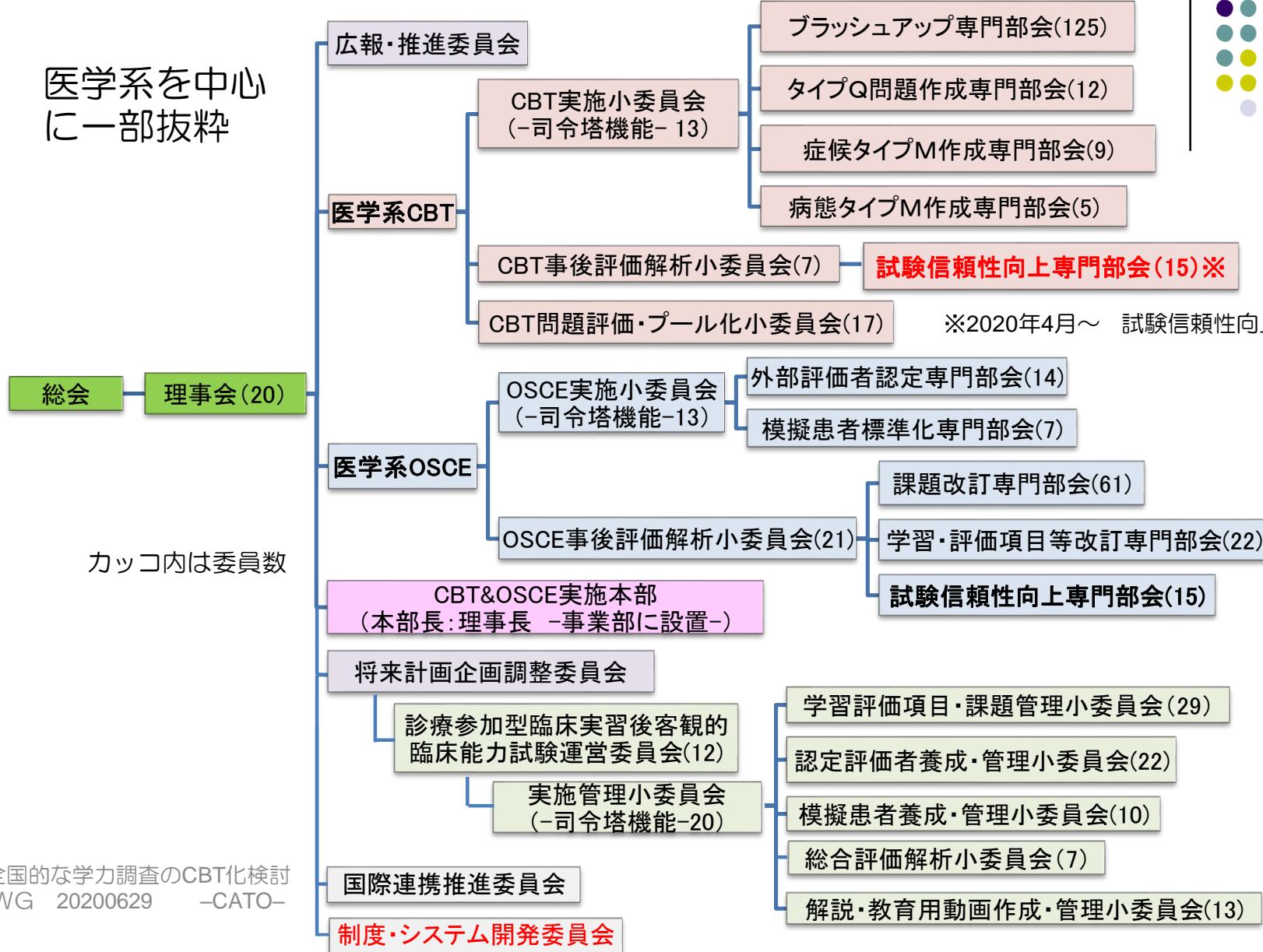
ブロックの構成

- ブロック1～4：各1時間：五選択肢択一形式で**60題**
⇒コア・カリキュラムA～D、FとE2、E3から出題
- ブロック5：1時間：多選択肢択一形式
⇒臨床症例問題：34題と基礎医学問題：6題の計**40題**
- ブロック6：1時間：順次解答4連問形式
⇒**10症例40題**
ブロック6については一度解答したら戻れない
- 全6時間で**320題**

共用試験の運営を支える組織



医学系を中心
に一部抜粋



CBT問題作成から成績返却まで



問題入力システム、問題作成テンプレート、問題作成マニュアル、ブラッシュアップチェックリスト等収載

- 各医科大学・医学部への作成依頼（計画的依頼、問題作成キット送付）
- 各医科大学・医学部でのブラッシュアップ

2018実績 221名参加
各大学に3名の推薦を依頼

- 各医科大学・医学部選出教員による中央ブラッシュアップ
- 問題評価・プール化小委員会による試行問題決定

中央ブラッシュアップの前後で機構委員によるPreブラッシュアップ、Postブラッシュアップを実施

2018実績 75回開催

- 試行問題+プール問題（採点対象）の組合せによる出題
- 問題セット作成と調整作業、総覧（実際に受験生が見るモードで）
- CBT実施キット準備・配布

2018実績 本試験129名、追再試験70名派遣

- 各医科大学・医学部でCBT実施（機構派遣監督者派遣）

- 各医科大学・医学部からCBT実施キット回収
- 機構内で自動採点（IRT標準スコア値推定も含む）

- 事後評価（試行問題+プール問題）

- 各医科大学・医学部へ成績返却
- CBT解析結果の公表



CBT問題の質の保証(1)

➤ 異なる時期に試験を行い、能力を評価する

- ✓ 問題プールから受験生ごとに異なる問題をランダムに出題する
- ✓ 出題セットの平均難易度に差がないようにする必要がある
- ✓ **項目反応理論** (IRT : Item Response Theory)
 - ① 試行（採点対象外）として出題した問題の評価（**200名以上**の受験生からの項目特性曲線：難易度と識別度）
 - ② 受験生ごとに異なる出題セット間の難易度調整（難易度の等質化）
 - ③ 受験生の能力評価 (IRT標準スコア)

➤ 医師国家試験の場合

- ✓ 事前に問題の評価ができないために事後に評価する
- ✓ 採点除外問題
 - ① 正解が複数存在：複数正解
 - ② 必修問題：「問題としては適切であるが、必修問題としては不適切」
→正解者は正解として扱い、不正解者は削除
 - ③ 一般問題：「選択肢の表現が不十分で正解を得ることができない」「設問が不適切で正解が得られない」「選択肢に誤りがあり正解が得られない」
 - ④ 採点除外問題：4題（113回）、3題（112回）、7題（111回）

CBT問題の質の保証(2)



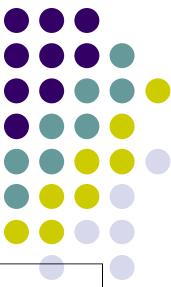
➤ 共用試験CBTの質保証

- ✓ 良質な問題の作成→全医学部に問題作成依頼（説明会の開催）
- ✓ 良質な問題のプール
 - ① 医学部委嘱教員による機構でのブラッシュアップ（5~6人のグループで2日間で50~60題）・・・2日間を3セット実施。各チームは必ず専門分野以外の教員も含める構成としている。
 - ② 機構内「問題評価・プール化委員会」でのブラッシュアップ
 - ③ 採点対象外として出題問題の事後評価（正答率、受験生からの審議依頼問題の評価）
- ✓ 採点問題のプールとセットとしての評価
 - ① 出題問題の特性変化の検討→大幅に正答率が変動した問題の削除（復元本により問題の漏洩？）
 - ② 問題の記載法の変更に対応→症例（外国人名、検査値など）の記載法
 - ③ 受験生から審議依頼問題の検討→トライアルだけでなく採点評価問題として出題した問題にも実施
 - ④ 出題セット作成前に出題可能問題の確認
 - ⑤ セット組の問題の再度の確認→実施小委員による総覧を実施
- ✓ **試験信頼性向上委員会**
 - ① CBTによる評価の妥当性を検討
 - ② 2か月に1回開催（CBT・教育測定学・統計学・テスト理論の専門家による検討）
- 以上のステップにより良質で評価可能な問題が出題されている

共用試験のCBTの内容



コアカリ項目	出題割合
A 基本事項 1 医の原則、2 医療における安全性確保、3 コミュニケーションとチーム医療、4 課題探求・解決と学修の在り方	約 4.2%
B 医学一般 1 個体の構成と機能、2 個体の反応、3 病因と病態	約 20.8%
C 人体各器官の正常・病態等 1 血液・造血器・リンパ系、2 神経系、3 皮膚系、4 循環器系、5 循環器系・・・14 耳鼻・咽喉・口腔系、15 精神系	約 37.5%
D 全身に及ぶ生理と病態 1 感染症、2 腫瘍、3 免疫・アレルギー疾患、・・・7 人の死、8 死と法	約 20.8%
E 診療の基本 1 症候・病態からのアプローチ、2 基本的診療知識、3 基本的診療技能	約 8.3%
F 医学・医療と社会	約 8.3%



国内外のCBT普及状況

国内外	用途		試験名称	内容	実施主体等
国内	学校教育において利用	入学者選抜	GTEC CBT	英語能力を測定	株式会社ベネッセコーポレーション、一般社団法人進学基準研究機構
			TEAP CBT	英語能力を測定	公益財団法人日本英語検定協会
		学力調査	PISA	国際的学習到達度調査	OECD
	専門教育において利用	医学・歯学共用試験CBT		診療参加型臨床実習に参加するための知識を問う	公益社団法人医療系大学間共用試験実施評価機構
		薬学共用試験CBT		同上	特定非営利法人薬学共用試験センター
		獣医学共用試験CBT		同上	特定非営利活動法人獣医学教育支援機構
	その他資格試験等で利用 (代表的なもののみ掲載)	ITパスポート試験		ITに関する知識を問う	独立行政法人情報処理推進機構
		SPI3などの就職時の適正テスト試験			
		漢検CBT		漢字能力を測定する技能検定	公益財団法人日本漢字能力検定協会
		英検CBT		英語能力を測定	公益財団法人 日本英語検定協会
国外	学校教育において利用	入学者選抜	Joint Entrance Examination	工学系大学に入学を希望する者に対して実施	インド
			GRE(Graduate Record Examinations)	学術系大学院に入学を希望する者に対して実施	アメリカ、カナダ
			ACT(American College Testing Program)	大学進学希望者に対して実施(CBT化を検討中)	アメリカ
			北京市英語リスニング試験	大学進学希望者に対して実施	中国北京市
		学力調査	PARCC(Partnership for Assessment of Readiness for College and Careers)	小学校3年から高校生を対象とした学力試験	アメリカ
			全国学力調査	中学1年生を対象とした試験	フランス
	専門教育において利用	USMLE(United States Medical Licensing Examination)		医師資格を得るための試験	アメリカ
		Florida Teacher Certification Examinations		教員資格試験	アメリカフロリダ州
		Common Admission Test		ビジネススクール入学試験	インド
	その他資格試験等で利用	全国計算機等級考試(National Computer Rank Examination)		コンピュータスキルを測定する試験	中国



米国の医師国家試験における CBT利用状況

USMLE(United States Medical Licensing Examination)

試験区分		概要等	CBT 利用	出題形式	受験手續先	受験料	日本での受験	
STEP1	基礎医学	医学に関する基礎知識を問う。出題科目は解剖学、生理学、生化学、薬理学、病理学、行動科学	○	全308設問出題(1ブロック1時間で44設問解答、7ブロック)	ECFMG	\$940	○	
STEP2	基礎臨床医学	CK	Clinical Knowledge: 臨床知識を問う。主に内科学、外科学、小児科学、精神医学、産婦人科学、公衆衛生、家庭医学、救急医学、生物統計学、医学文献等	○	最大355設問出題(1ブロック1時間で最大45設問解答、8ブロック) マルチメディア対応(心臓の音を聞かせたり、患者さんの動画を見させて症状を選択肢から選択等)	ECFMG	\$940	○
		CS	Clinical Skills: 臨床技能を測る。評価項目は3つ。 Communication and Interpersonal Skills (CIS)、Spoken English Proficiency (SEP)、Integrated Clinical Encounter (ICE)	×	試験時間8時間。12人の模擬患者 (Standardized Patient)と相対します。1人の患者役毎に15分、その後10分間でPatient Noteを作成し、これを12回繰り返す	ECFMG	\$1,580	×
STEP3	総合的臨床プロセスとケースシミュレーション	1日目	Foundations of Independent Practice (FIP)	○	全233設問出題(1ブロック1時間で38～40設問解答、6ブロック)	FSMB		×
		2日目	Advanced Clinical Medicine (ACM)	○	全180設問出題(1ブロック45分で30設問解答、6ブロック)	FSMB		×
			Computer-based Case Simulations(CSS)	○	受験生1人あたり患者13人分のケースについて、患者1人につき約10分から20分かけて解答。シミュレーションテスト。	FSMB	\$875	×

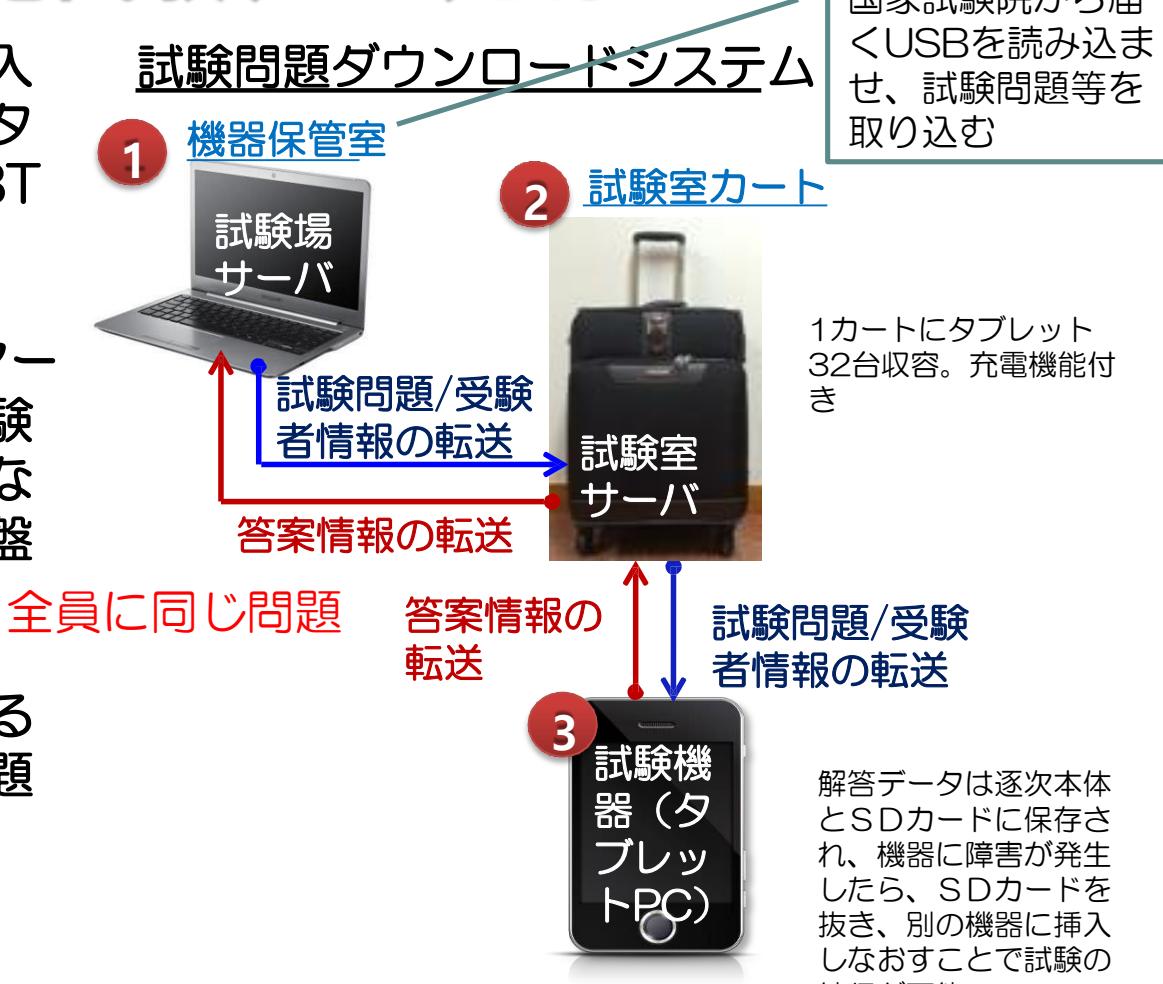
受験料は2019年度



韓国の国家試験の事例：Smart Device-based Test (SBT)

コンピュータ化試験(SBT)とは

- 保健医療人国家試験に導入推進中であるコンピュータ化試験（2020年からPBTから移行予定）
- タブレットPCなどのスマートデバイスを活かし、試験の進行、採点、成績管理などが可能なオフライン基盤の試験方式
- 動画、音声などが含まれるマルチメディア問題が出題可能

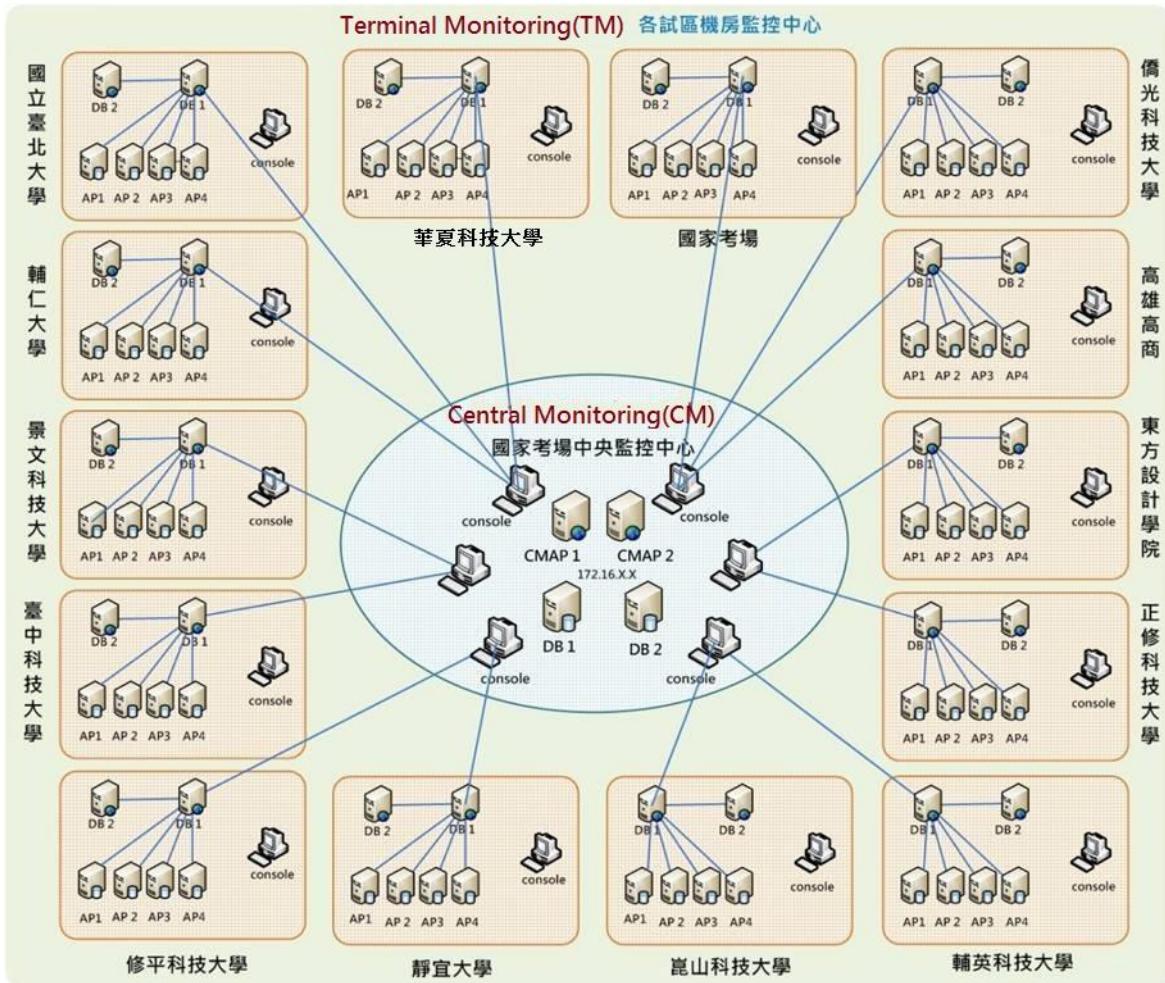


試験時は1回3万Wで学生に貸出

受験料は筆記試験28.7万W、実技62万W

運営コストはPBTと変わらないため、受験料値上げの予定なし

台湾の国家試験の構成



【私見】学力調査においては、

①各学校に中間サーバをおかなくとも、都道府県単位で中間サーバを置き、各教育委員会ごとに仮想サーバを用意してもよいのではないか。その仮想サーバでは学力調査の実施だけにとどまらず、分析・事後評価での利用、共用コンテンツ管理機能、臨時のアンケート実施機能など付加機能を考えてもよいのではないか。

(その際は、仮想サーバのメンテナンスは各都道府県等で行う)

②地域特性、規模特性に応じて、韓国のようなS B T方式、USB方式、場合によっては障害者差別解消法対応でP B Tを認めるなどハイブリッド対応があつてもよいのではないか。



CBTの工夫(1)

- CCS (computer-based case simulation)
 - ✓ USMLEステップ3に導入
 - ✓ コンピュータに救急症例を提示
 - ✓ 患者への対応をコンピュータに入力
 - ✓ 適切な対応を臨床推論する
 - ✓ 受験時間は症例により異なる（10～20分）
 - ✓ 受験症例は9例から13例に増加
- 順次回答形式 (computer-patient) シリーズ問題
 - ✓ 診察手順に従ったMCQ（臨床推論）
 - ✓ 医療面接、身体診察、臨床検査、診断＆病態生理
 - ✓ 世界で共用試験だけ



CBTの工夫(2)

- マルチメディアを活用したMCQ
 - ✓ 聴診での音声を利用したMCQ
 - ✓ キーフィルムだけでなく医療現場と同様に複数の画像から受験生が選択して解答
 - ✓ 患者・医師の動画からの出題
 - ✓ 診察手順に従ったMCQ

！重要なのは、どのような能力を測定したいのか。
その次にコンピュータの機能をどう活用するか。



大動脈弁閉鎖不全症での心音

38歳の男性。労作時呼吸困難と下腿浮腫とを主訴に来院した。25歳ころから高血圧と心雜音とを指摘されていた。3年前から山道や階段を上がるときに軽度の呼吸困難を感じ、立ち止まるようになった。2年前から食後や重いものを持った際に胸部の重苦しい感じが出現した。3か月前から平地歩行でも呼吸困難を自覚するようになり、両下腿に浮腫が出現している。
既往歴と家族歴とに特記すべきことはない。喫煙歴と飲酒歴とはない。

聴取される心雜音はどれか。

- a 連續性雜音
- b 心膜摩擦音
- c 収縮期クリック音
- d 拡張期ランブル
- e to-and-fro 雜音

正解 e

■患者の心臓聴診所見を音声で示す。





意識レベル；Japan Coma Scale

35歳の女性。患者の様子がおかしいと家族に連れられて来院した。

現病歴：3日前から38°C台の発熱があり、市販の総合感冒薬を内服していた。昨日は朝から頭痛を訴え、終日臥床していた。今朝、家族が訪室したところ、呼びかけに対する反応がおかしいことから受診した。

既往歴・家族歴：特記すべきことはない。

現 症：体温 38.6°C。脈拍 96/分、整。血圧 132/80 mmHg。呼吸数 18/分。SpO₂ 99% (room air)。

■搬入時の診察を動画で示す。

意識レベルは Japan Coma Scale でどれか。

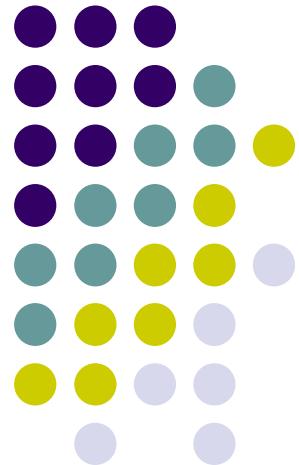
- a I -1
- b I -2
- c I -3
- d II -10
- e II -20

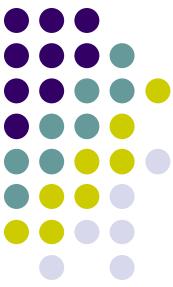
正解 c



CBT導入の概要等

- ・トライアルの概要
- ・経費負担、人的負担





共通の評価システムとしての共用試験

「臨床実習に臨む学生の能力・適正について、全国的に一定の水準を確保するとともに、学生の学習意欲を喚起する観点から、

共通の評価システムを作る

ことを検討すべきである。」

1999年2月「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」
(21世紀医学・医療懇談会第4次報告)

とても紙ではできない



委託事業者に丸投げせず、問題作成から試験実施まで自分たちで
PDCAサイクルを廻す

1. 最初のハードル

- 1) 全大学共通の評価基準・評価問題に基づく、全大学共用の試験を検討する。これまでやったことがない事業に対する各大学の不安、特に feasibility の観点から。

→ 可能な限りのコンピュータ化を検討（19年前）

- 2) 共用試験構想の周知活動を始めるも、「受ける日も違う、受ける問題も違う、そんな状況で公平・公正な合否判定ができるのか」という強い意見。

→ 諸外国の同様の試験の状況を調べ、CBTとIRTをセットで導入する必要があることを確認

1999年 USMLE STEP1 CBTを開始 (with IRT)



CBT&IRTをセットで

1. CBTとIRTをセットで導入することを前提に改めて必要な要件を検討

1) 試験システムの根幹

- 非同時実施
- 非同一問題
- 追試験、再試験
- セキュリティ
- ブループリント（モデル・コア・カリキュラムに基づいた出題）
- IRT適用（適切な露出回数の確保、プール問題の確保）

2) 学生の適切な能力評価のために必要な配慮

- 問題形式（出題形式）
- マルチメディア対応
- アダプティブ対応
- CBTに慣れていない学生への対応（公開問題、体験テスト、操作説明ツール）

3) フィードバック(受験料を徴収して行う試験、ハイステークスな試験の観点から)

- 受験生へのフィードバック
- 実施大学へのフィードバック
- 問題作成、事後評価、実施管理、プール問題管理部会等へのフィードバック
- 行政へのフィードバック

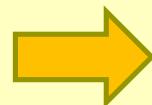
CBT&IRTをセットで -2-



1. CBTとIRTをセットで導入することを前提に改めて必要な要件を検討

4) feasibility

- ・実施大学の環境（ICTインフラ、ICT要員）
- ・運用手順の確立、障害対応
- ・各種マニュアル、手引き（問題作成キット、事前動作確認キット、試験実施キット）
- ・受験票の設計
- ・受験環境の確認
- ・受験生、実施大学責任者からの実施後調査票
- ・機構派遣監督者によるモニタリングシステム

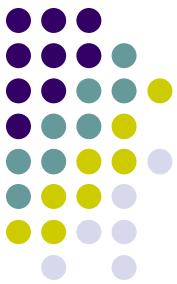


必要な要件の検討を踏まえて、トライアル実施計画を策定並行して、

- ①トライアルサポート施設（※）の設置要請
- ②トライアル実施のための財政支援を要請
- ③正式実施後の運営母体となる社団法人設立の準備を開始。

共用試験の意義を再確認と周知。

トライアルの経過



1. トライアル0（ゼロ）

1) CBT導入可能性の検討のためのプロトタイプシステム開発

- 100人に100問の問題をコンピュータ出題するシステムの構築（非ランダム出題）

2. トライアル1（80大学中80大学で実施 受験者数7,915人）

1) 100人以上の受験生に100問コンピュータ出題するシステムの構築

- 2ブロック制
- タイプAのみ
- ランダム出題（受験生毎に異なる出題）
- 難易度調整なし
- 素点評価

2) 問題作成システムの開発、各大学にシステムを配布、問題作成用PCを1大学1台配布

3) 出題問題管理システムの開発

4) 試験実施システムの開発

トライアルの経過 -2-



3. トライアル2 (80大学中80大学で実施 受験者数7,827人 追再試験30大学 518人)

1) 100人以上の受験生に300問コンピュータ出題するシステムの構築

- 6ブロック制
- タイプA、W、Q（順次解答型連問形式）
- ランダム出題（受験生毎に異なる出題）アンカー問題20%
- 難易度調整なし
- 素点評価

2) 追・再試験実施機能の追加

3) 出題形式の多様化対応（順次回答2連問、4連問）

4) 事前体験ツールの作成（体験テスト、操作説明ツール）

5) 試験開始許可キーの導入

6) フライング設定機能の付加

7) パニックID機能の付加

8) 事後評価システムの開発

9) IRT適用準備

トライアルの経過 -3-



4. トライアル3（80大学中80大学で実施 受験者数8,109人 追再試験38大学 701人）

1) 100人以上の受験生に320問コンピュータ出題するシステムの構築

- 6ブロック制
- タイプA、W、Q（順次解答型連問形式）、R、L（多選択肢連問形式）
- ランダム出題（受験生毎に異なる出題）アンカー問題20%
- 素点による難易度調整
- 素点評価

2) プール問題と試行問題の区分出題機能の付加

3) 出題形式の多様化対応（多選択肢形式）

4) 重複問題排除機能の付加

5) 視覚素材の出典管理機能付加

6) 高解像度ディスプレイ対応

7) 事後評価機能の強化（経年データ表示、累積統計）

8) 問題作成システム評価エディションの開発（修正履歴、管理情報、ブラッシュアップ機能強化）

トライアルの経過 -4-



5. 最終トライアル -正式実施を想定した運用のチェック-

(80大学中80大学で実施 受験者数7,836人 追再試験41大学 612人)

1) 100人以上の受験生に320問コンピュータ出題するシステムの構築

- 6ブロック制プラス1アンケートブロック
- タイプA、Q（順次解答型連問形式）、L（多選択肢連問形式）、タイプW・R廃止
- ランダム出題（受験生毎に異なる出題）アンカー問題20%
- IRTによる難易度調整
- 素点評価、IRT評価

2) 受験者固有番号の追加

3) 各大学フィードバックデータの定型化（個人別成績表、結果総表、年度総表）

4) アンケートブロックの導入

5) サイトサーバにおけるリアルタイムモニタ機能追加

6) 試験時障害対応機能の追加（スタンドアローンモード、予備端末移動、予備受験票対応 障害時トレース機能）

7) ブロック解答中、ブロック間休憩時のロック機能の追加

8) 試験実施システム事前動作確認機能の追加



トライアル期間を通じたプール問題の蓄積状況

1. トライアル1

80大学からタイプA（5肢択一形式） 9,919設問提出、
トライアルを通じて、1万問を超える
プール問題（採点対象問題）を確保
ブラックアップ委員会による精選作業で、2,791設問採択（採択率28.1%）
事後評価解析委員会による試験実施後の問題評価により 2,305設問をプール問題候補とした。

2. トライアル2

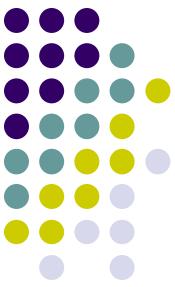
80大学からタイプA（5肢択一形式） 9,322設問提出
ブラックアップ委員会による精選作業で、3,723設問採択（採択率40.0%）
事後評価解析委員会による試験実施後の問題評価により 2,908設問をプール問題候補とした。

3. トライアル3

80大学からタイプA（5肢択一形式） 7,108設問提出
ブラックアップ委員会による精選作業で、3,803設問採択（採択率53.5%）
新形式問題作成チームにより、タイプW、R、Qについて合計449設問作成。
事後評価解析委員会による試験実施後の問題評価により 3,875設問をプール問題候補とした。

4. 最終トライアル

80大学からタイプA（5肢択一形式） 3,175設問提出
タイプL（多選択肢2連問形式・・・鑑別診断問題） 2,080設問
タイプQ（順次解答4連問形式・・・臨床推論問題） 1,964設問
ブラックアップ委員会による精選作業で、問題タイプ別にそれぞれ、
1,316設問（採択率 41.4%）、712設問（採択率 34.2%）、412設問（採択率21.0%）採択
実施小委員会による新規作成、旧問題サルベージにより、タイプA259設問、タイプQ56設問作成
事後評価解析委員会による試験実施後の問題評価により 3,172設問をプール問題候補とした。



トライアル期間を通じたプール問題の蓄積

- アイテムバンクを作成し、それに基づき出題
- 毎年参加大学の教員が多くの試験問題を作成することから、他の試験に比べると新作の問題数が多く、良質のアイテムバンクの構築と多数の新陳代謝が可能
- 新作問題の評価にはトライアルを4年間実施し、十分な問題数を確保
- 本試験ではアイテムバンクから抽出した問題と新作問題を同時に出題して新作問題の項目特性値の評価を行う

トライアルのテストデザイン

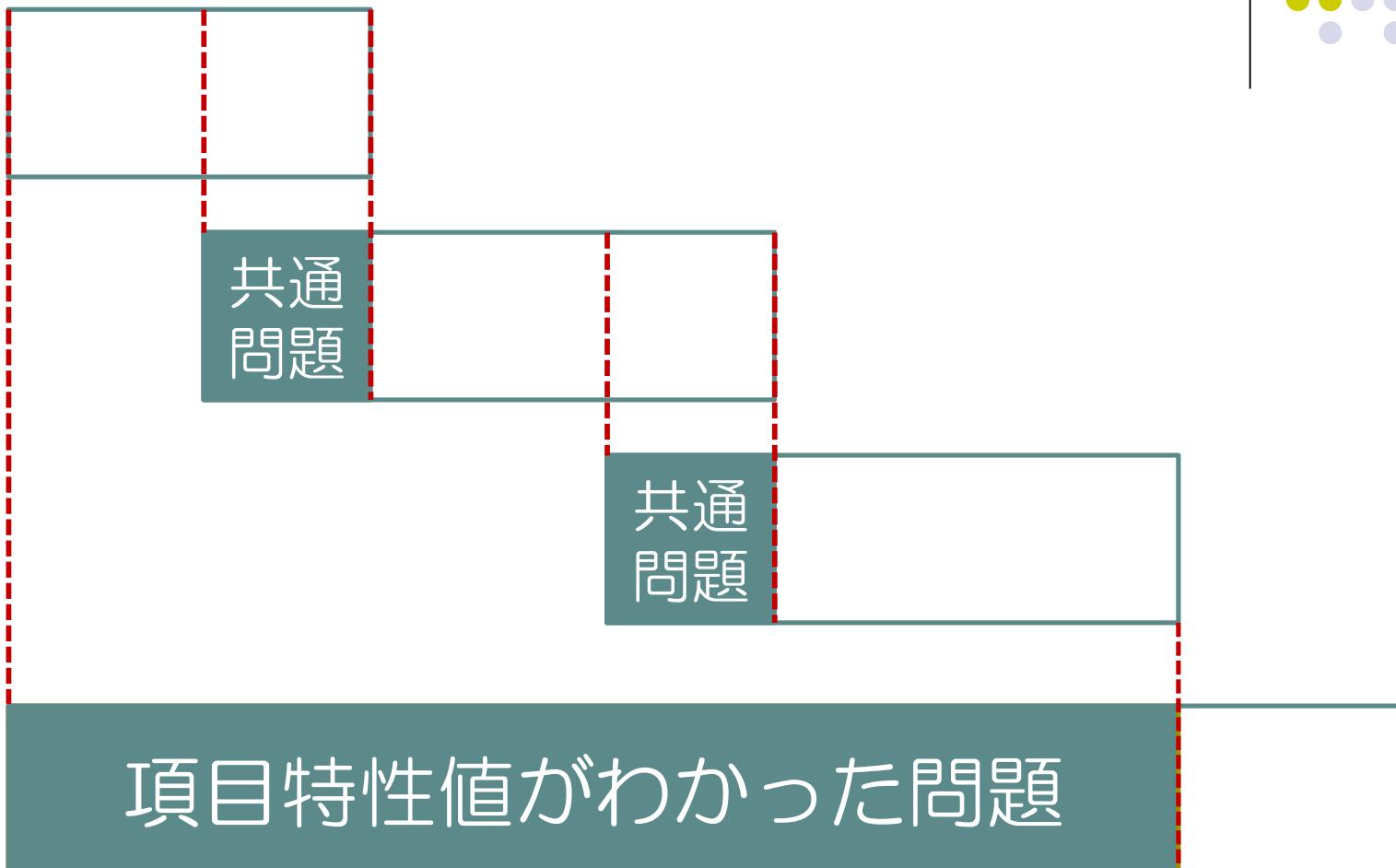


Trial
1

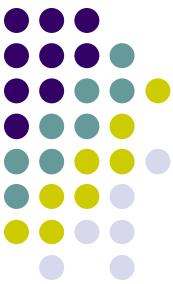
Trial
2

Trial
3

Trial
4



新作問題とアイテムバンク



アイテムバンク

2006年

アイテムバンク抽出問題

新作

2007年

アイテムバンク抽出問題

新作

2008年

アイテムバンク抽出問題

新作





トライアル時（4年間）の経費・人的負担

1. 経費負担（文部科学省関係：207,097千円）

1) 科学研究費補助金（37,300千円）

- 2001年～2002年度科学研究費補助金（特別研究促進費）

「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評価研究」

- 2003年～2004年年度科学研究費補助金（特別研究促進費）

「共用試験正式実施に向けての医歯学教育システム開発と学習到達度評価法確立の調査研究」

2) 大学改革推進等補助金

3) ソフト開発経費、導入経費（119,609千円）

2. 経費負担（任意団体共用試験実施機構会費収入（各大学からの拠出）（382,023千円）

3. 人的負担（支援：東京医科歯科大学に医歯学教育システム研究センター（MDセンター）を設置し、専任教職員配置）-社団法人設立後、一部人員転籍）

※1：トライアル期間中は

- 受験料は徴収していない
- 各大学医学部の教職員に対して謝金等は支払われていない。

※2：上記の経費はCBTに加えて、技能試験であるOSCEの経費も含んだものである。



正式実施以降の経費負担

1. CBT&OSCEをセットで受験料を設定

本試験：25,000円（2005年から2014年まで28,000円）

追試験：無料

再試験：25,000円（2005年から2014年まで28,000円）

※トライアルの実績から総事業費を算出し、受験予定者数で除して受験料を算出

総事業費 302,422千円 想定受験数 医学系8,000名、歯学系3,000名 計11,000名

CBT関連事業 127,444千円

OSCE関連事業 80,218千円

調査研究・開発事業 46,000千円

研修・出版・広報事業 38,760千円

国際協力事業 10,000千円

2. 教員に対する謝金

1日 10,000円

3. 機構の業務を委嘱する各大学教職員、研修指定病院医師、その他関係者の移動時の傷害に対する傷害保険料を機構にて負担



試験の運用体制 -CBTの場合-

1. 試験の問題管理、出題管理、実施管理等運営管理全般

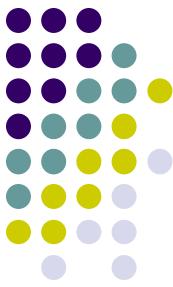
- ・実施小委員会、事後評価委員会、問題評価・プール化小委員会、機構で対応
- ・試験実施時に機構事業部に実施本部を設置するとともに機構派遣監督者（※）を派遣する。

※：機構派遣監督者は 共用試験CBTの質、運営の公正性・公平性を担保し、さらなる向上を図るために、共用試験CBTが公正かつ適切に行われているかを確認し、試験全体を監督するとともに、試験が厳正かつ公正に実施されたかを最終的に判断します

2. 各実施大学での環境整備と試験実施時の体制

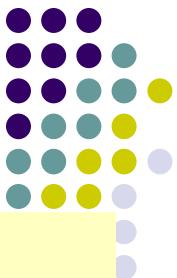
- ・CBTで使用するサイトサーバーおよびPCは各実施大学で準備。
- ・各実施大学の環境が適切かどうかを動作確認キットを使用して確認する。（動作確認データは機構に送付して、試験が実施できるかを機構でチェック）
- ・機構において、毎年、各実施大学の実施体制を調査。
　統括実施責任者、会場実施責任者、サイトマネージャー、連絡担当者等

3. 試験当日のシステムサポートセンターの開設（技術的な対応を業者に委託）、システムトラブル以外は機構に設置する実施本部で対応



共用試験で項目反応理論が用いられる背景

- 共用試験CBTは、同一時期に一斉試験が困難なため、異なる時期に、異なる場所で、異なる問題の試験を実施しても公平な評価が得られる試験方式を採用する必要があった。
- CBTでは、過去に出題した問題について、複数回の内容のチェックと解答状況の解析を行い、良質かつ適切と考えられる問題を継続的に蓄積した。
- CBT実施時には、このプールした問題バンクの中から、受験生ごとに異なった問題がランダムに抽出されて出題される。
- 異なる問題が出題されるための不公平感をなくすために、受験生一人一人に出題される問題セット間の難しさの差ができるだけ小さくなるように設定する。
- この方法だけでは、一人一人のテストの難しさの差を完全になくすことはできないため、受験生に出題された項目の特性を考慮して問題セット間の難しさの差に影響されない成績評価の方法として項目反応理論を用いることとなった。



IRTの適用 (CBTとセット)

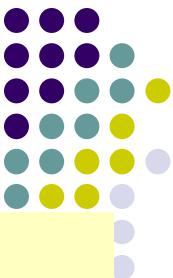
1. 項目（出題問題）の事後評価

<目的>

良質かつ適切（臨床実習開始前の難易度として）と考えられる項目の継続的な蓄積

<IRTの適用>

- 項目特性曲線のパターン（急峻なパターン→識別力の視覚化）
- 項目困難度（難しさのレベルの定量化）
- 項目識別力（識別力の定量化）
- 項目特性値の標準誤差（項目特性値の推定精度の定量化）
- 基本となる項目特性値と直近の年度で推定された項目特性値の比較（項目特性値の変化状況の把握）



IRTの適用 (CBTとセット)

2. 出題セットの作成

<目的>

ランダム出題のため異なる問題セットが受験生に出題されることからセット間の難易度を小さくする。

<IRTの適用>

- 項目困難度によるセット間の難易度差の調整



- 情報関数を用いたセット間の難易度差と識別力差の調整



IRTの適用 (CBTとセット)

3. 試験成績の評価とフィードバック

<目的>

受験生に出題された項目の特性を考慮して項目セット間の難しさの差に影響されない成績評価を行う。

<IRTの適用>

出題された問題セット間の難易度差に影響されない受験生の成績評価としてIRT標準スコアを用いる。

1) 全国的評価

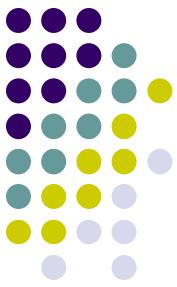
- ・結果表、冊子の提供：毎年の全体的成績、月別の推移、再試験受験者の本試験と再試験の比較、問題タイプ別比較、コアカリ大分類別比較など
- ・講演会：経年変化の提供

2) 個人の評価

- ・個人別成績表：IRT標準スコアの個人別成績と学内分布における位置の提供、コアカリ大項目別問題タイプ別レーダーチャートによる学内平均との比較

3) 学内評価

- ・総表（個人別の成績一覧表）：各大学における受験生評価に利用可能な資料の提供



IRTの適用 (OSCEとセット)

<目的>

OSCEを含むパフォーマンス評価は、評価者の差という誤差要因を含むため、この影響を除いた評価として、近年パフォーマンス評価に用いられてきている項目反応理論を用いる。

<IRTの適用>

- 評価者の配置方法の適切化を行い、評価者の特性を求める。
- 評価者の特性に影響されない受験生の評価を行う。
- 将来的には出題される課題の等化、認定評価者を用いた等化、ビデオを用いた等化を行い、全国的な評価を検討する。



共用試験の現状について
お話ししました。
ご清聴、
ありがとうございました。





参考資料

1. CBT問題例
2. その他



五肢択一問題

- Aタイプ（単純択一形式）
- 正解は唯一であるようにする。ここでいう「唯一」とは「唯一絶対」ということばかりでなく、「相対的に正しい」ことでも良い。（one-best）
- 90%と20、10、0%でも選択することが可能
- 例題：抗核抗体が陽性となるのはどれか。

A SLE	②
B 肝硬変	1
C 心筋梗塞	0
D 関節リウマチ	2
E ネフローゼ症候群	0

「適正度」（共用試験）



例1:A-1-(2)-3)

患者が自己決定できない場合の対処法を説明できる。

インフォームドコンセントについて正しいのはどれか。

- A. 新薬の臨床試験においては必要ない。
- B. 看護師が医師に代わって行っても良い。
- C. 未成年者の患者では、保護者のみに行えばよい。
- D. 進行癌の場合、病名をまず患者の男性の親族に伝える。
- E. 意識不明の身元不明患者では、医師の判断で救命ができる。

正解：E



例2：A-2-(1)-4)

医療の安全性に関する情報（薬害や医療過誤の事例、やってはいけないこと、模範事例等）を共有し、事後に役立てるための分析の重要性を説明できる。

救急患者の採血検体を検査部へ持っていくよう指示された人（立っている女性）が、検体を検査技師に手渡す際の写真（カラー表示）を示す。

写真に示された現行の方法について優先的に改善すべき点はどれか。



- A. 検査の項目をシートで説明する。
- B. 検査の項目を口頭で告げ、技師も復唱する。
- C. 患者の氏名を口頭で告げる。
- D. 採血管に患者の氏名を手書きで記入する。
- E. シートのID番号と氏名を検体に貼付する。

正解：E

70

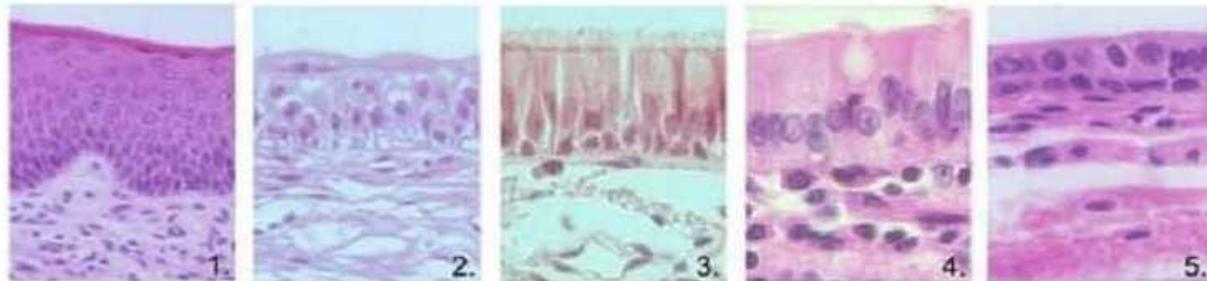


例3：B-1- (2) -①-1)

上皮組織と腺の構造と機能を説明できる。

膀胱や尿管の内面をおおう上皮はどれか（写真はカラー表示）。

- A. 1
- B. 2
- C. 3
- D. 4
- E. 5



正解：B

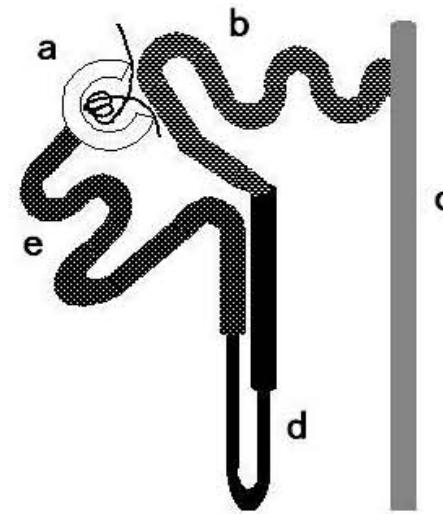


例4: C-8- (1) -3)

腎の機能の全体像やネフロン各部の構造と機能を概説できる。

腎ネフロンの模式図（白黒表示）を示す。
a～eの部位の機能で正しいのはどれか。

- A. アルブミンはaでほとんどろ過される。
- B. ブドウ糖の再吸収は主にbで行われる。
- C. 水分の再吸収量が最も多いのはcである。
- D. d周辺の間質の浸透圧は高張である。
- E. バソプレシンが作用するのは主にeである。



正解:D



例5：D-2- (1) -1)

腫瘍の定義を説明できる。

悪性腫瘍の一般的特徴でないのはどれか。

- A. 核異型
- B. 破壊性増殖
- C. 遠隔臓器転移
- D. 細胞極性の消失
- E. 周囲組織との境界が明瞭

正解：E



多選択肢連問問題

□ タイプM

- 選択肢が6つ以上（8～15肢が目安）である。
- 正解肢は1つ
- 共用試験CBTで使用されているが、国家試験でも採用されるようになった。
- 選択肢数が多いので、単純想起だけでなく、**解釈力を測定できる**。
- 設問文によっては**応用力や問題解決力**も測定できる。
- 偶然正解となる可能性は激減する。

CBTでは

症候から診断を問う（鑑別診断）⇒「症候EMI」（F領域）36設問
ある病態・事象の原因を問う⇒「病態EMI」（C, E領域）4設問



症候タイプM：「発熱」をきたす疾患

選択肢

- A 亜急性甲状腺炎
- B 悪性リンパ腫
- C 感染性心内膜炎
- D 胸膜炎
- E 結核
- F 成人Still（スチル）病
- G 全身性エリテマトーデス
- H 多発性動脈炎
- I. 膿胸
- J. 肺炎
- L 麻疹
- M リウマチ熱



症候タイプM「発熱」（1/2）

54歳の女性。2週持続する発熱のために来院した。以前に健康診査で心雜音を指摘されている。1か月前に歯周病の治療を受けている。身長155cm、体重50kg。体温37.7 °C。脈拍88/分、整。血圧110/48mmHg。皮疹は認めない。頸部に圧痛はなく、腫大したリンパ節も触知しない。胸骨左縁第3肋間を最強点とする拡張期雜音 (Levine 2/6) を聴取する。呼吸音に異常を認めない。血液学所見：白血球14,000 (桿状核好中球 20%、分葉核好中球55%、好酸球2%、好塩基球 1%、単球 3%、リンパ球19%)、血小板 20万。血液生化学所見：AST 20 IU/L、ALT 24 IU/L、LDH 250 IU/L (基準176~353)。CRP 12.4mg/dL。血液培養で*Streptococcus viridans*を認める。

考えられるのはどれか。

（正解：C 感染性心内膜炎）

選択肢

- A 亜急性甲状腺炎
- B 悪性リンパ腫
- C 感染性心内膜炎
- D 胸膜炎
- E 結核
- F 成人Still（スチル）病
- G 全身性エリテマトーデス
- H 多発性動脈炎
- I 膽胸
- J 肺炎
- L 麻疹
- M リウマチ熱



症候タイプM 「発熱」 (2/2)

22歳の女性。2か月前から37℃台の微熱が続き、全身倦怠感が強いため来院した。身長158cm、体重48kg。体温37.4℃。脈拍80/分、整。血圧100/80mmHg。血圧に左右差を認めない。顔面と上肢に皮疹を認める。頸部に圧痛はなく、腫大したリンパ節も触知しない。心音と呼吸音に異常を認めない。左右ともに遠位指節間関節と中手指節関節の腫脹と圧痛を認める。尿所見：蛋白（2+）、糖（-）、潜血（+）。血液学所見：赤血球330万、Hb 10.2g/dL、白血球3,400、血小板8.5万。免疫学所見：抗核抗体640倍（基準40未満）、抗Sm抗体陽性。

考えられるのはど�か。

(正解： M リウマチ熱)

選択肢

- A 亜急性甲状腺炎
- B 悪性リンパ腫
- C 感染性心内膜炎
- D 胸膜炎
- E 結核
- F 成人Still（スチル）病
- G 全身性エリテマトーデス
- H 多発性動脈炎
- I. 膽胸
- J. 肺炎
- L. 麻疹
- M リウマチ熱



病態タイプM 「病態と細胞像」

選択肢

- A Bリンパ球
- B Tリンパ球
- C 形質細胞
- D 好酸球
- E 好塩基球
- F 好中球
- G 樹状細胞
- H 肥満細胞
- I マクロファージ



病態タイプM 「病態と細胞像」 (1/2)

18歳の男子。2時間前から息苦しくなり、ゼイゼイして目が覚め、救急外来を受診した。呼吸困難は増悪している。小児期にアトピー性皮膚炎の既往があり、季節の変わり目に時々同様な症状があった。チアノーゼは認めず、呼気延長と吸気時の喘鳴を聴取する。呼吸機能検査で1秒率の低下を認める。

喀痰中に多く認められる細胞はどれか。

(正解:D 好酸球)

選択肢

- A Bリンパ球
- B Tリンパ球
- C 形質細胞
- D 好酸球
- E 好塩基球
- F 好中球
- G 樹状細胞
- H 肥満細胞
- I マクロファージ



病態タイプM 「病態と細胞像」 (2/2)

48歳の男性。昼食でカニを食べた15分後から、かゆみの強い発疹が出現し、次第に地図状に拡大してきたため来院した。気管支喘息と花粉症の既往がある。呼吸数 18/分。血圧 118/74mmHg。皮膚描記法で陽性である。

この患者の病態に最も関係しているのは
どれか。

(正解:H 肥満細胞)

選択肢

- A Bリンパ球
- B Tリンパ球
- C 形質細胞
- D 好酸球
- E 好塩基球
- F 好中球
- G 樹状細胞
- H 肥満細胞
- I マクロファージ



順次解答4連問形式問題(タイプQ)

Q=quadruple (4部分からなる)

- 臨床推論能力
- Paper patientによる診察
- 基本パターン
 - 第1設問 医療面接
 - 第2設問 身体診察
 - 第3設問 検査
 - 第4設問 この症例の病態生理
- コンピュータの特性を利用して
前に戻れない



ブロック6 順次解答4連問 (60分で10症例)

- 問題形式は**5肢択一**、連問形式で4問出題される
- 第1問→第2問→第3問→第4問と順次解答する
- 臨床推論能力をみる**症例問題**とする
「後戻りできない」という特徴あり
 - 第1設問：原則、**医療面接**で確認すべきこと
 - 第2設問：**身体診察**でとくに重要なこと
 - 第3設問：診断確定に必要な**検査**に関するこ
 - 第4設問：症例の**病態生理**や**生命科学**に関するこ



第1設問：医療面接

52歳の女性。午前4時頃に右肋骨の下の方が重苦しく、時々差し込むような痛みを感じて目覚めた。痛みがだんだんひどくなってきたため午前5時に救急外来を受診した。吐き気はあるが吐いてはいない。小太りである。

この患者にまず聞くことはどれか。

- A 血尿があるか。
- B 黒色便があるか。
- C 残尿感があるか。
- D 昨夜、脂っこい食事を摂ったか。
- E 家族にも同じ症状の人がいるか。

(正解:D)



第2設問：身体診察

52歳の女性。午前4時頃に右肋骨下の重圧感と差し込むような 痛みを感じて目覚め、痛みが増強したため午前5時に救急外来 を受診した。恶心はあるが、嘔吐はない。
昨夜は中華料理を食べた。 身長 152cm、体重 65kg。
体温 37.8°C。脈拍 76/分、整。血圧 124/78 mmHg
。心音と呼吸音に異常は認めない。腹部は平坦で、肝・脾を触知しない。

予想される身体所はどれか。

- A 金属性の腸雜音
- B 肋骨脊柱角叩打痛
- C Murphy (マーフィー徵候)
- D Blumberg (ブルンベルグ) 徵候
- E McBurney (マックバーニー) 徵候

(正解：C)



第3設問：検査

52歳の女性。右肋骨下の重圧感と差し込むような痛みを感じて目覚め、痛みが増強したため救急外来を受診した。昨夜は中華料理を食べた。

身長 152cm、体重 65kg。体温 37.8°C。脈拍 76/分、整。血圧 124/78mmHg。腹部は平坦で、肝・脾を触知しない。右肋骨下領域に圧痛があり、軽く叩打すると痛みが増強する。Murphy (マーフィー)

徵候を認める。 血液学所見：赤血球 510万、Hb 14.5g/dL、Ht 46%、白血球 14,000、血小板 18万。血液生化学所見：総ビリルビン 1.8mg/dL、直接ビリルビン 1.5mg/dL、AST 38 IU/L、ALT 37 IU/L、アミラーゼ 80 IU/L（基準37~160）。

まず行う検査はどれか。

- A 腹部単純CT
- B 腹部超音波検査
- C 腹部エックス線撮影
- D 上部消化管造影
- E 上部消化管内視鏡検査

(正解：B)



第4設問：病態生理

52歳の女性。右肋骨下の重圧感と差し込むような痛みを感じて目覚め、痛みが増強するため救急外来を受診した。昨夜は中華料理を食べた。

身長 152cm、体重 65kg。体温 37.8°C。脈拍 76/分、整。血圧 124/78mmHg。腹部は平坦で、肝・脾を触知しない。右肋骨下領域に圧痛があり、軽く叩打すると痛みが増強する。Murphy (マーフィー) 徴候を認める。血液学所見：赤血球 510万、Hb 14.5g/dL、Ht 46%、白血球 14,000、血小板 18万。血液生化学所見：総ビリルビン 1.8mg/dL、直接ビリルビン 1.5mg/dL、AST 38 IU/L、ALT 37 IU/L、アミラーゼ 80 IU/L（基準37~160）。腹部超音波像（白黒表示）を示す。

この患者で脂っこいものを食べた後に起こった腹痛の発生機序に関わるホルモンはどれか。

- A モチリン
- B ガストリン
- C セクレチン
- D ソマトスタチン
- E コレシストキニン

(正解：E)





項目反応理論(CATO冊子より)

① 項目反応理論について：

「項目（Item）」とは、試験を構成する1つ1つの問題のこと（以下、個々の試験問題を項目と表記する）、「反応（Response）」とは、その項目に正答するか誤答するかの状況を表現するものです。項目反応理論とは、項目の特性（難易度、識別力）が判明している場合、その項目に対する反応（解答状況）を用いて、当該試験の結果から測定できる能力を推測するものです。大規模試験の項目作成・実施・評価・運用のための優れた実践モデルとして世界的に定着しています。

② 共用試験の特殊性：

「同一問題を用いて、同一時期に一斉に試験をすること」が公平であるとの考えがありますが、各大学のカリキュラム（臨床実習開始時期）が異なることから、共用試験の全国一斉実施は困難です。そのため、異なった時期に試験を実施しても公平な評価が得られる試験方式が求められ、項目反応理論を用いた試験方式が適切な方法として採用されました。

③ 項目反応理論による成績評価：

基本となるのは能力値です。能力値は、項目反応理論で測定しようとする能力を表す尺度であり、 θ （シータ）と表記します。受験生個人の推測される能力を表すもので、全受験生では平均値0、標準偏差1の標準正規分布を仮定します。また、項目の特性を示す曲線の横軸として、能力の尺度を表しています。この能力値という物差しを意味のある数値とするために、基準となる集団を決めて【医学系（2012～2014年度共用試験）正式実施結果、歯学系（2013年度共用試験）正式実施結果】、その集団のデータが標準正規分布（平均値500、標準偏差100）するように計算を行っています。この基準を固定することにより、現在のデータとの相対的位置が分かり、経年変化を見ることが可能となります。基準集団の見直しは定期的に行う必要があります。医学系は基準集団の見直しを行い、第11回（2016年度）共用試験から（2012～2014年度共用試験）正式実施結果を用いています。



項目反応理論(CATO冊子より)

① 項目反応理論について：

「項目（Item）」とは、試験を構成する1つ1つの問題のこと（以下、個々の試験問題を項目と表記する）、「反応（Response）」とは、その項目に正答するか誤答するかの状況を表現するものです。項目反応理論とは、項目の特性（難易度、識別力）が判明している場合、その項目に対する反応（解答状況）を用いて、当該試験の結果から測定できる能力を推測するものです。大規模試験の項目作成・実施・評価・運用のための優れた実践モデルとして世界的に定着しています。

② 共用試験の特殊性：

「同一問題を用いて、同一時期に一斉に試験をすること」が公平であるとの考えがありますが、各大学のカリキュラム（臨床実習開始時期）が異なることから、共用試験の全国一斉実施は困難です。そのため、異なった時期に試験を実施しても公平な評価が得られる試験方式が求められ、項目反応理論を用いた試験方式が適切な方法として採用されました。

③ 項目反応理論による成績評価：

基本となるのは能力値です。能力値は、項目反応理論で測定しようとする能力を表す尺度であり、 θ （シータ）と表記します。受験生個人の推測される能力を表すもので、全受験生では平均値0、標準偏差1の標準正規分布を仮定します。また、項目の特性を示す曲線の横軸として、能力の尺度を表しています。この能力値という物差しを意味のある数値とするために、基準となる集団を決めて【医学系（2012～2014年度共用試験）正式実施結果、歯学系（2013年度共用試験）正式実施結果】、その集団のデータが標準正規分布（平均値500、標準偏差100）するように計算を行っています。この基準を固定することにより、現在のデータとの相対的位置が分かり、経年変化を見ることが可能となります。基準集団の見直しは定期的に行う必要があります。医学系は基準集団の見直しを行い、第11回（2016年度）共用試験から（2012～2014年度共用試験）正式実施結果を用いています。



項目反応理論(CATO冊子より)

④ 共用試験における成績評価：

IRT標準スコアを用いています（これは機構独自の名称で、2012年度から名称を変更しました）。このスコアは「能力値(θ) $\times 100 + 500$ 」の式で求めています。偏差値と同様に基準集団での相対的な位置を示すもので、その年度の受験集団内の位置ではありません。

⑤ 項目特性曲線、項目特性値：

項目反応理論における重要な要素の項目特性曲線は、項目の特性を曲線で表すものです（次頁の図）。横軸は、前述した能力値(θ)です。縦軸は、各能力値(θ)の当該項目に正答する確率となります。この曲線は、始めは緩やかに増加し、徐々に傾きが大きくなり、ある点から後は、増加率が鈍化し、やがて正答する確率が1に近づきます。その曲線の変わり目（変曲点）の位置を決めるのが b （項目困難度）であり、そこで増加率は最大となります。また、そこでの傾きを示すのが a （項目識別力）となります。この曲線の位置と傾きにより、当該項目の特性が視覚的に分かります。

⑥ 項目困難度：

項目特性曲線から求められる項目特性値は、いくつかの種類がありますが、この特性値は項目特性曲線の位置を表します。共用試験が利用している項目特性曲線では、正答する確率が0.5の時の項目特性曲線と交わる位置の能力値(θ)で表します。一般的には b と表記しています。 b が大きいということは能力値(θ)が高くないと正答できないということであり、項目の難しさを表現しています。

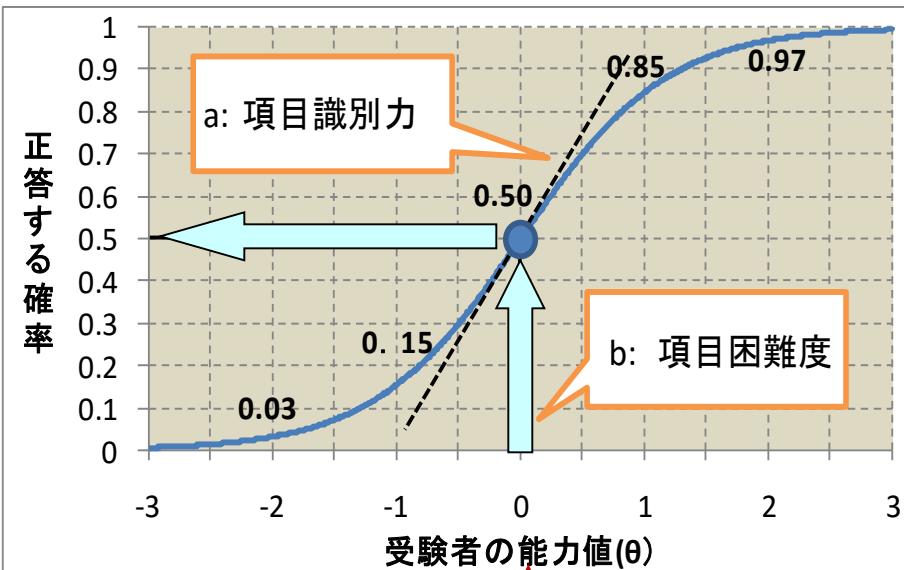
⑦ 項目識別力：

この項目特性値は、項目特性曲線の能力値(θ)が $= b$ における傾きに比例し、一般的には a と表記しています。 a が大きいということは、 b を挟む能力値(θ)の前後で正誤が比較的はっきり見分けやすくなります。

項目反応理論



項目特性曲線



各大学に問題作成

↓
中央のブラッシュアップ委員会で修正等
↓ (近年の採択率≈80%)

新規問題としてプレテストに出題 (採点対象外)

↓
プレテストに出題された個々の試験問題 (項目)
について、学生の能力 (総合得点) とその問題を
正答する確率の関係を、2パラメータ成長曲線を
用いて回帰分析し、パラメータ a, b を求める。

$$\text{正答する確率} = P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))}$$

↓
事後評価で、正答率、識別指数、点双列相関係数
とともに両パラメータ値を基準として適正な問題
を選んでプール問題とする

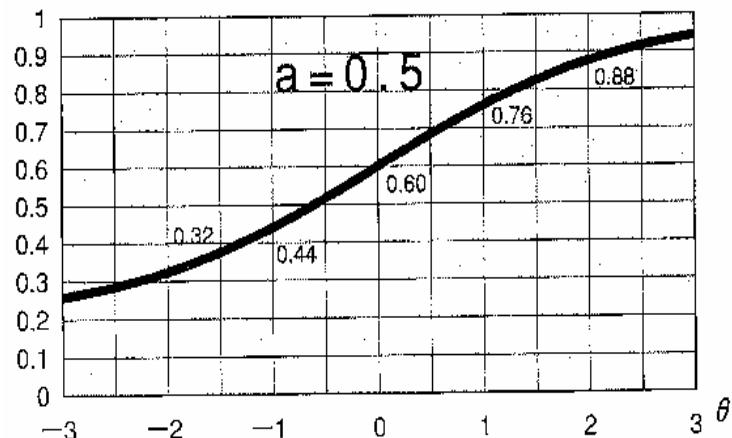
項目特性値 (項目識別力と項目難易度)



a 項目識別力

能力差のつく問題かどうかの指標

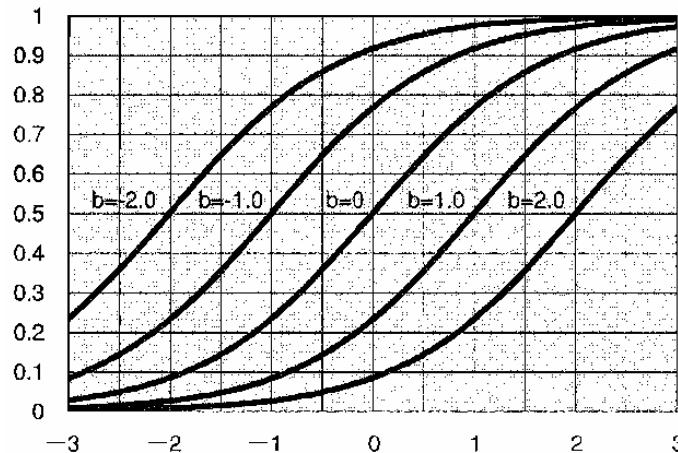
1. 特性曲線の最大の傾きを表す。
2. a が大きいほど識別力が大きい
3. 通常は0.3~2.0程度



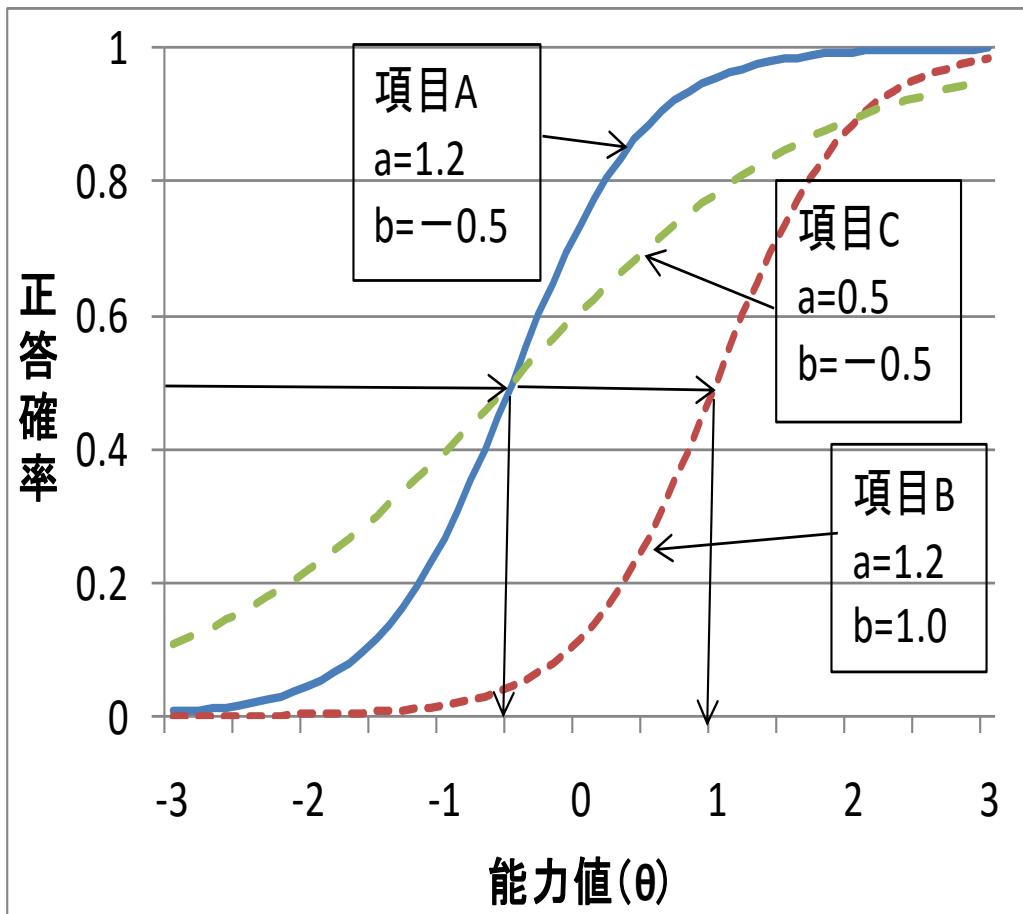
b 項目難易度

問題の難しさを表す指標

1. その問題を正解できる確率が0.5の能力
2. b が大きいほど難しい問題
3. 多くは-2~+2の範囲程度



項目特性の評価(3項目の例)



- 項目Aは能力の低い段階で正答する割合が上昇(易しい項目)
- 項目Bは能力の高い人でないと正答できない(難しい項目)
- 項目Cは能力の低い人でも正答可能で上昇が緩やか(識別力の低い項目)

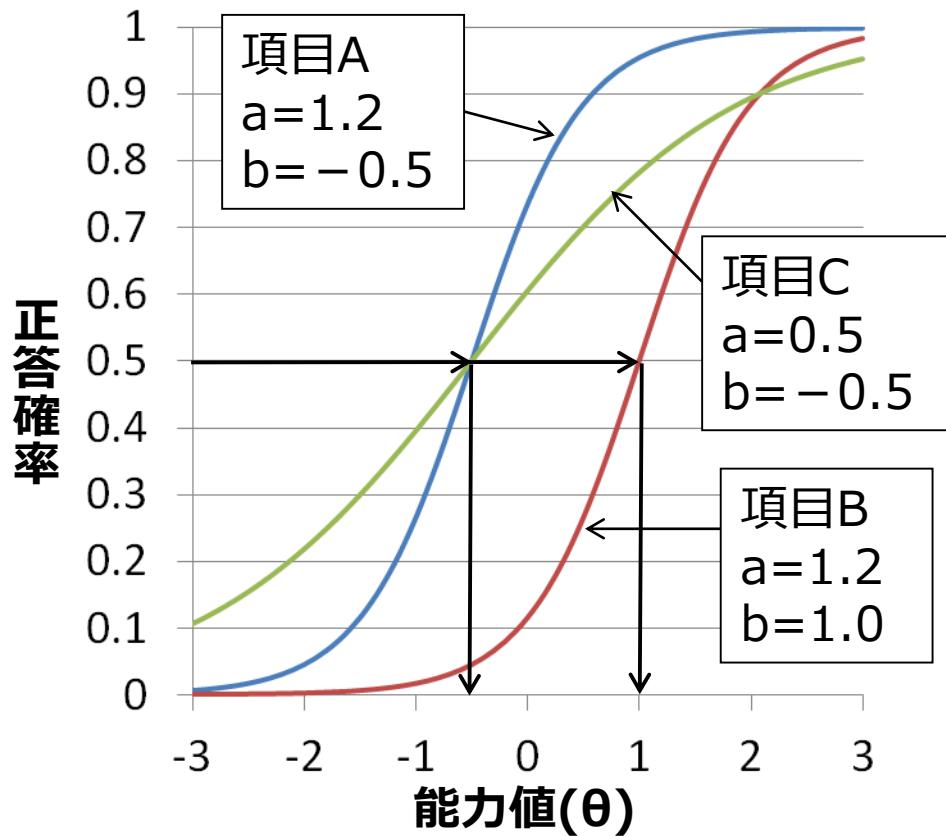
項目特性の評価(3項目の例)



項目A：能力の低い段階で正答する割合が上昇（易しい項目）

項目B：能力の高い人でないと正答できない（難しい項目）

項目C：能力の低い人でも正答可能で上昇が緩やか（識別力の低い項目）



「従来のテスト」と「項目反応理論」の 違い



	従来のテスト	項目反応理論
能力の表現方法	正答数, 正答率など	問題の難易度・識別度からの統計値で表現
項目の難易度	経験に基づく	事前テストのデータに基づく
事前テスト	不要	問題の難易度・識別力推定に必要
異なるテスト間での比較	不可能	可能