

共用試験CBTにおける 項目反応理論の適用と実践

公益社団法人
医療系大学間共用試験実施評価機構
仁田 善雄

1

講演内容

1. 共用試験の成立までの流れ
2. 共用試験とは
3. 共用試験CBTの構造
4. 共用試験への項目反応理論の適用
5. 項目反応理論を用いた良質なプール問題の選択
6. 項目反応理論を用いた共用試験の結果と効果
7. まとめ

2

共用試験の成立までの流れ

3

医・歯学教育改革

1996年文部省有識者懇談会報告書

- 知識の伝授のみで、態度・技術の習得が充分行われていない。
- 技術の教育とその評価の必要性の指摘

同 第4次報告

- 臨床実習の充実が必要 → 診療参加型臨床実習
- そのための適切な進級認定システムの構築と進路指導の充実が必要

4

診療参加型臨床実習とは

医・歯学生

- 指導医の監督のもと
- 診療チームの一員
- 実際の患者診療に従事

※一定の医行為が許容されると同時に患者診療への責任が生じる。

5

臨床実習の充実のため 必要とされる医行為

社会通念上医学生に許容される 医行為の実施条件

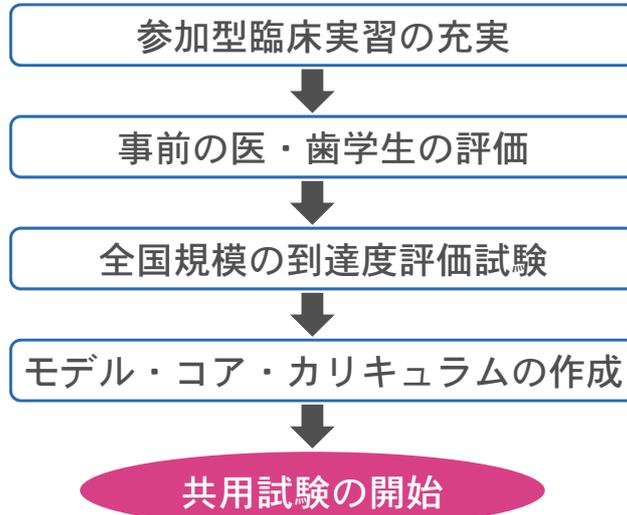
- 侵襲性の高くない、一定のものに限る
- 指導医によるきめ細かな指導・監視のもとに行われる
- 事前に医学生の評価を行う
- 患者などの同意を得て実施する

厚生省臨床実習検討委員会（前川委員長）最終報告 1991年

* 現在見直しが考えられている

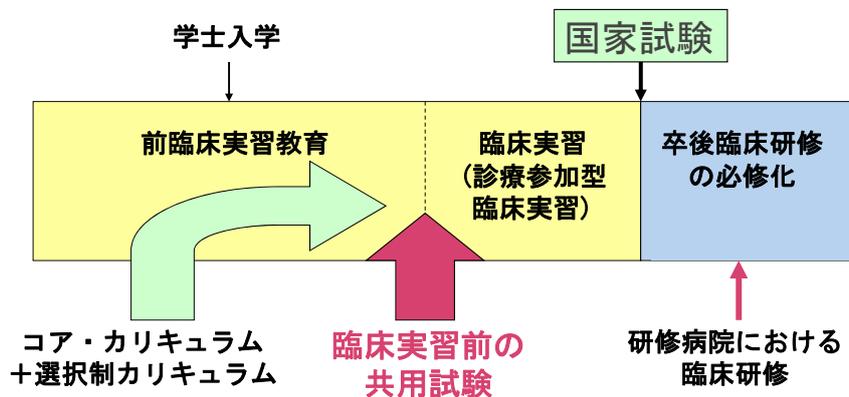
6

共用試験の成立までの考え方



7

医学教育の構造



8

平成28年度版医学教育 モデル・コア・カリキュラム

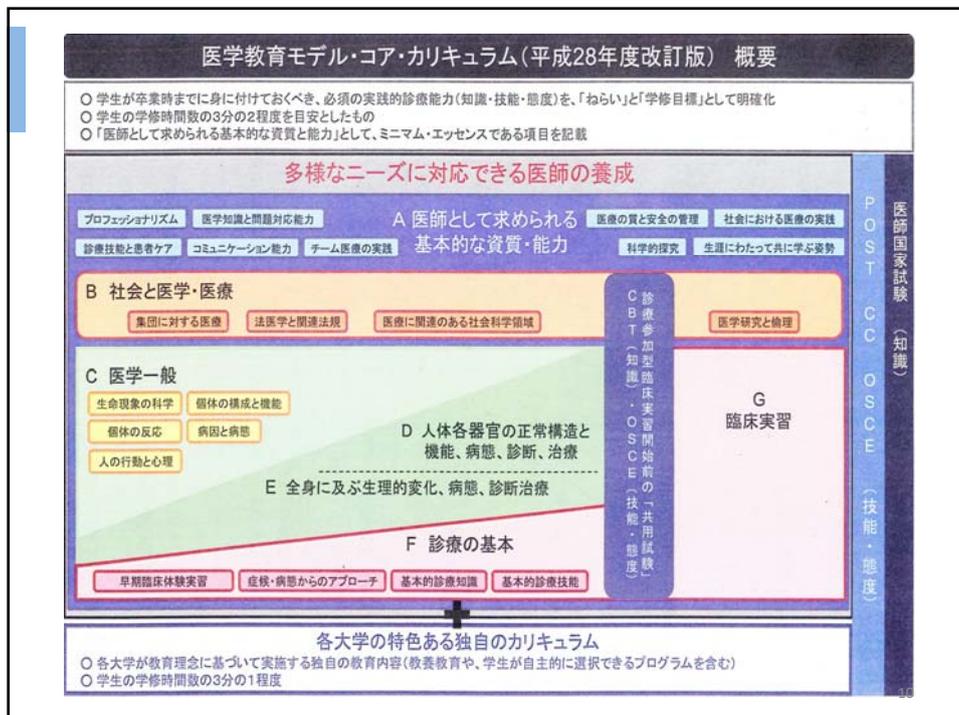
基本理念と背景

- ① 多様なニーズに対応できる医師の養成
- ② 社会の変遷への対応
- ③ 卒前・卒後の一貫性
- ④ 医学・歯学における「基本的資質・能力」の共有

* 各大学の医学教育のうち、学修時間数の約2/3を目安とした内容・分量について、体系的に整理された全国共通のカリキュラム

- * A 「医師として求められる基本的な資質・能力」
- B 「社会と医学・医療」
 - C 「医学一般」
 - D 「人体各器官の正常構造と機能、病態、診断、治療」
 - E 「全身に及ぶ生理的変化、病態、診断、治療」
 - F 「診療の基本」
 - G 「臨床実習」

9



平成28年度版歯学教育 モデル・コア・カリキュラム

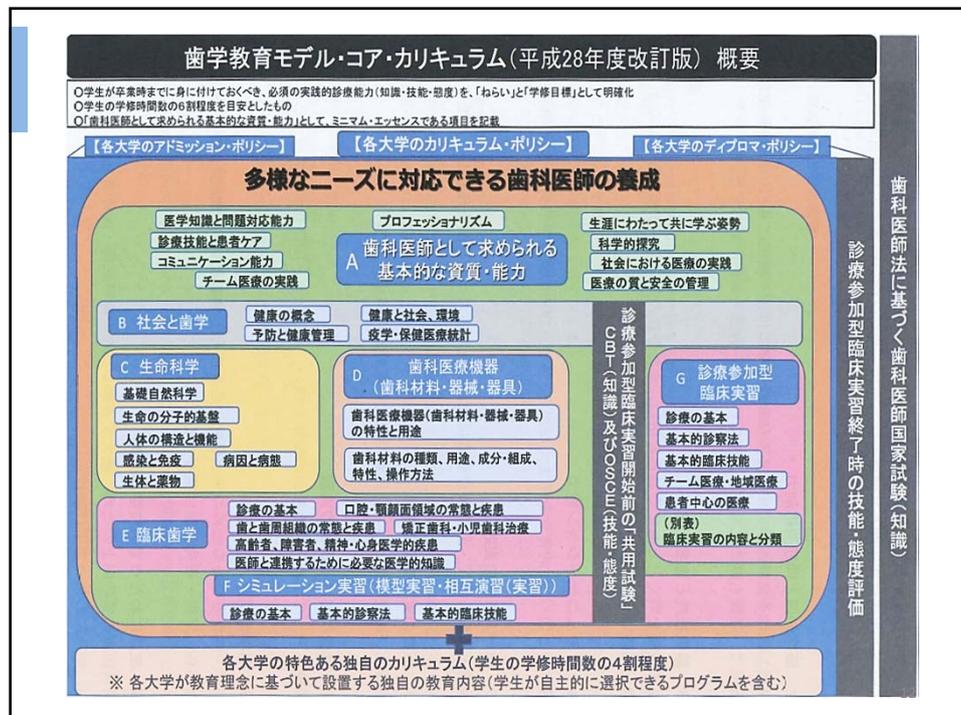
基本理念と背景

- ①多様なニーズに対応できる歯科医師の養成
- ②社会の変遷への対応
- ③医療提供体制の地域包括ケアシステムにおけるチーム医療・多職種連携
- ④卒前・卒後の一貫性
- ⑤医学・歯学における「基本的資質・能力」の共有
- ⑥高等教育における質保証

* 各大学の歯学教育のうち、学修時間数の6割程度を目安とした内容・分量について、体系的に整理された全国共通のカリキュラム

- * A 「歯科医師として求められる基本的な資質・能力」
 B 「社会と歯学」
 C 「生命科学」
 D 「歯科医療機器（歯科材料・器械・器具）」
 E 「臨床歯学」
 F 「シミュレーション実習（模型実習・相互実習（実習）」
 G 「臨床実習」

11



共用試験とは

13

試験の内容

CBT × OSCE

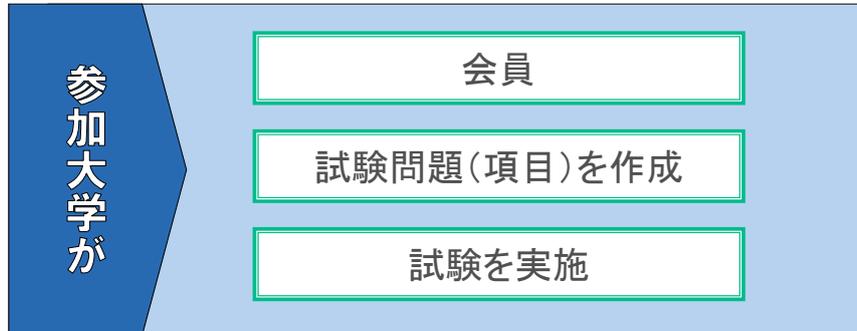
Computer Based Testing Objective Structured Clinical Examination

知識の総合的理解力 基本的診療技能・態度

目的: 臨床実習開始前の到達度評価
臨床実習において必要となる知識, 技能, 態度を評価

14

共用試験の基本的考え方



- 試験問題(項目)作成者と異なる人による項目評価・成績評価
- Peer review

15

共用試験実施運用組織

- 平成14年～17年までは医学系80大学、歯学系28大学が自主的に参加した任意団体「共用試験実施機構」と「医歯学教育システム研究センター(全国共同利用施設)」が協力して試験を実施運用した。
- 平成17年3月に「社団法人医療系大学間共用試験実施評価機構(現在、公益社団法人)」を設立。医学系80大学(現在、82大学)、歯学系28大学(現在、29大学)の社員としての参加により以後の正式実施を運用している。

16

大学間の共用試験システム

効率的な試験問題の作成と実施

各大学が共同で質の高い総合試験問題を作成、実施する。

共通の評価基準

臨床実習に参加している医学生・歯学生がある一定の能力を有する学生であることを社会に対して保証する。

大学間の流動性を高める

学外病院での診療参加型臨床実習の拡充を促す。

17

共用試験のこれまでの流れと基本的考え方

- 2002年～2005年の4年間にわたり、トライアルを実施
- 2005年12月から正式実施が開始、現在に至る
- 学生はCBTとOSCEの両方の合格が臨床実習への進級に求められる
- 共用試験の合格レベルは実施大学が決定
- 医学部長病院長会議が設定したCBTの推奨最低合格ラインが存在する（歯学系も現在検討中）
- 大学が合否判定を行い、医学部長病院長会議にチューデントドクター証の発行を依頼する

18

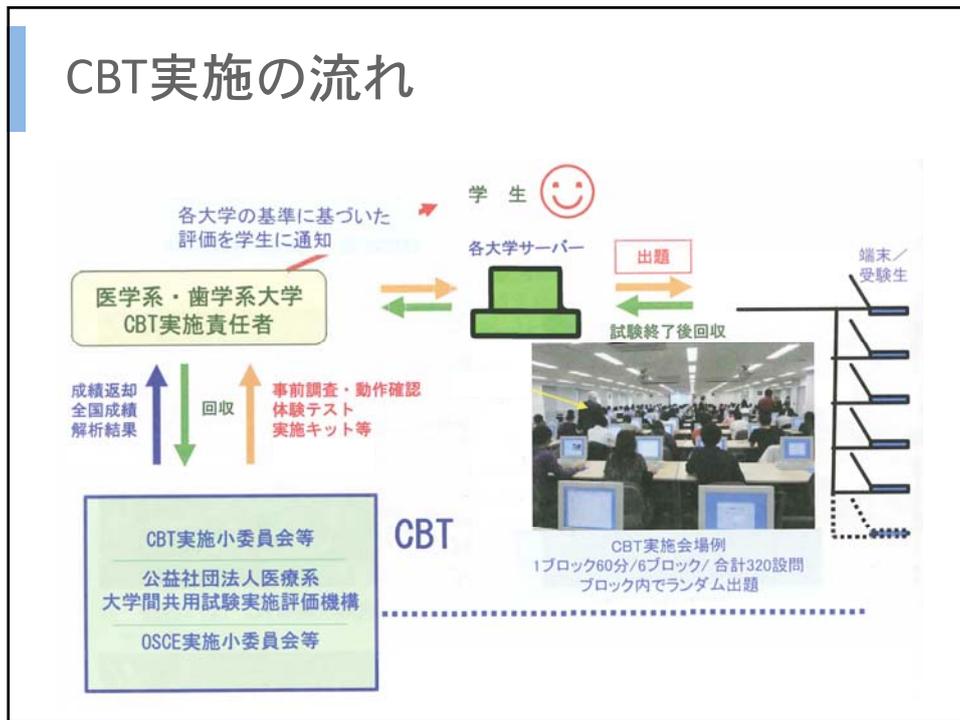
共用試験CBTの構造

19

CBTの構造

- 試験実施
- 問題作成
- 問題評価
- 問題セットの作成
- フィードバック

20



試験実施：試験システムの事前確認ツール

試験システムの動作確認用のツールの配布

各大学のネットワークシステムの動作を事前に検証する

試験システムの事前体験可能なツールの配布

- 受験生に試験システムの事前体験してもらう
- 同時に大学の試験運営のトライアルにもなる
- 受験者と実施主体の両方に十分な事前体験が可能となる

試験実施：機構派遣監督者制度

- 試験の信頼性の観点から他大学の監督者を機構から派遣（機構派遣監督者）
- 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

23

試験実施：役割分担

参加大学	機構
<ul style="list-style-type: none">• 試験会場の決定• 設営• システムの事前検証• 試験監督• 運用の実施	<ul style="list-style-type: none">• 試験システムの信頼性の検証と改善• 適切な試験セットの作成• 試験監督者の派遣（機構派遣監督者）• 試験実施本部の運用• 試験成績の作成• 大学へのフィードバック

- ◆ 大学により試験環境やコンピュータシステムが異なるため，システムの対応に費用が発生する

24

現在のCBT

医学系【6ブロック、320設問、6時間】

- ・タイプA：五選択肢択一（240設問）
- ・タイプM：多選択肢択一（40設問）
- ・タイプQ：順次解答4連問五選択肢択一
（10症例40設問）

歯学系【6ブロック、320設問、6時間】

- ・タイプA：五選択肢択一（240設問）
- ・タイプL：多選択肢順次解答2連問（20設問）
- ・タイプW：順次解答2連問（20設問）
- ・タイプQ：順次解答4連問五選択肢択一（40設問）

25

試験問題作成の流れ

機構

問題作成の依頼(不足領域など)
問題作成システムの提供



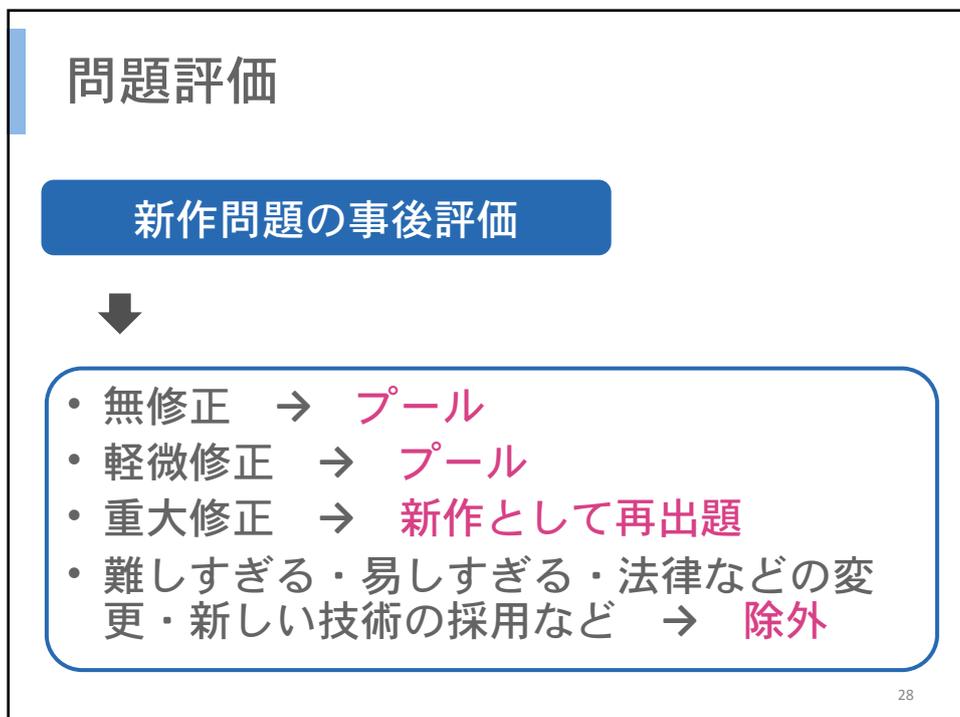
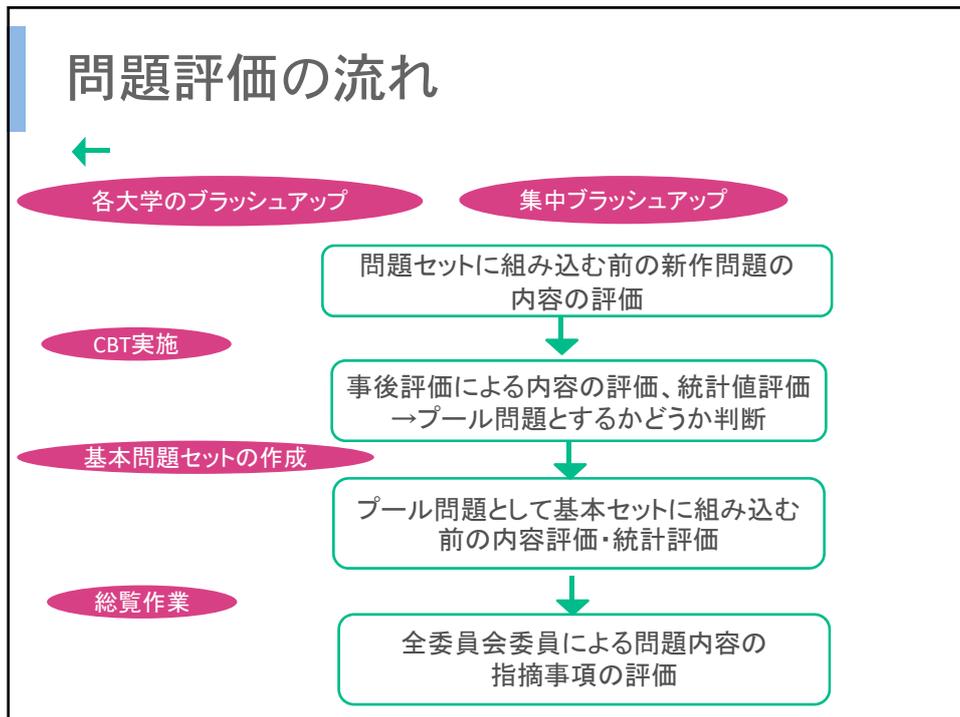
大学

問題作成、ブラッシュアップ、CD郵送



機構

全大学から1, 2名の各専門領域の教員を集め、
集中ブラッシュアップ、実施小委員会委員の最終
ブラッシュアップ、プール化委員の新作として
適切な露出数などの検討も踏まえた最終確認



試験問題のセット作成

問題評価・プール化委員会による基本となるプール問題による試験問題セットの作成



全委員会委員による総覧作業



総覧作業の結果の確認と基本セットの変更作業

29

成績などのフィードバック

参加大学

- 成績返却
- 試験データの解析結果の提供
- テスト理論について専門家による講演会開催
- テスト理論の説明文書・用語説明・パンフレットの提供

受験生

- 個人成績表の提供
- 試験全体の成績などを網羅した冊子の提供

30

個人成績表

①「個人成績表」の返却フォーマット
各大学のCBT実施キット等が機械に全て返却された後、10日以内に各大学宛発送します。
2013年度から正確な表現とするために正規分布の数値に近い値に表記を変更しました。

G-02
取扱注意

GATO 共用試験2018医学系CBT個人成績表

試験実施ID: M18-0001-00000FAD1 (採点日: 2018/08/22 発行日: 2018/08/22)
実施日: 2018/08/20 学校名: ○○○○大学医学部
受験番号: MA001 試験会場: ○○○○大学医学部
試験種別: 本試 ○の教室
学籍番号: 17000001 本試験受験者数: 100 (受験予定数: 100 欠席者数: 0)
氏名: 受験者○○○ 備考:

正答率: 75.40 あなたのCBT標準スコアの順位は100人中29位です。
CBT標準スコア: 550
6段階評価: 4

(あなたの成績は誰より高かったか)
基準集団の10%以下(分布範囲に位置する)
1 基準集団の2.5%未満
2 基準集団の2.5%以上～4.5%未満の範囲内
3 基準集団の4.5%以上～15.0%未満の範囲内
4 基準集団の15.0%以上～50.0%未満の範囲内
5 基準集団の50.0%以上～84.1%未満の範囲内
6 基準集団の84.1%以上～97.5%未満の範囲内
7 基準集団の97.5%以上

【コアカリ別・問題形式別集計】

コアカリ別	正答率	大学平均	大学最良値	大学中央値	大学最悪値	大学SD
大問						
小問						
選択						
記述						
英検						
英語						
数学						
物理						
化学						
生物						
倫理						
法律						
経済						
政治						
社会						
総合						
合計						

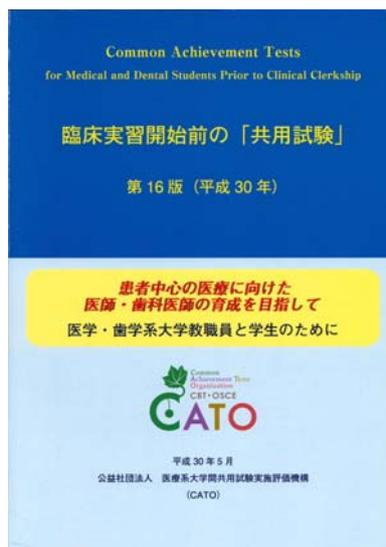
【あなたの大学の正答率分布】

【コアカリ別・問題形式別得点状況】

Copyright © 2018 公益社団法人 医療系大学間共用試験実施評価機構 . All Rights Reserved.

31

冊子の配布 (全学生ならびに教員)



32

共用試験への 項目反応理論の適用

33

試験方式

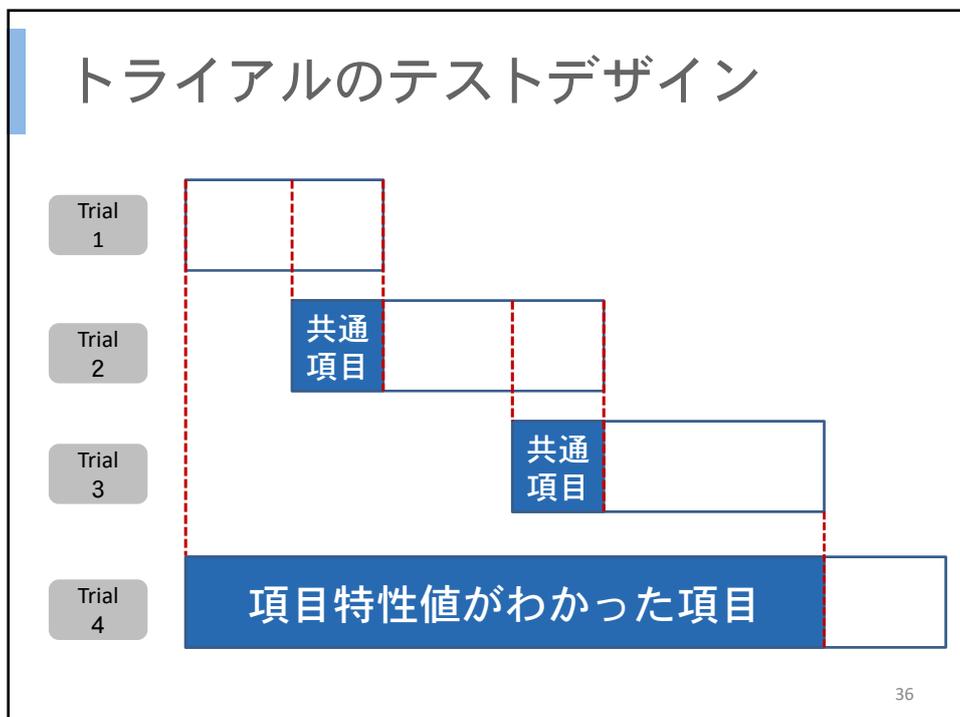
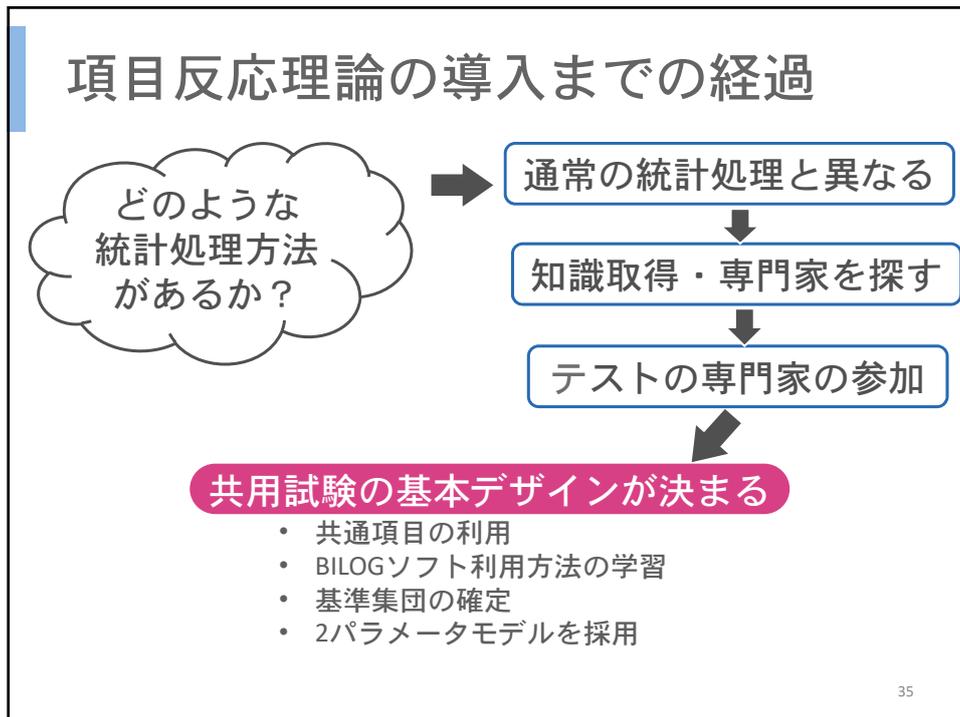
臨床実習開始時期が大学により異なる

同一時期での単一フォームの試験が不可能

受験生ごとに異なる問題セットで行う試験

- ◆ 再試験も行うため、同一個人に異なる問題セットが出題されるようにシステム制御を行っている
- ◆ 項目反応理論を用いて、項目評価、成績評価を行っている。

34



項目プールと項目バンク

項目バンク:
テストを構成するための
項目群を保持している

項目プールの一部分を
なす

項目プール:
項目特性値が既知
(項目識別力、項目困難度
の統計値が得られている)
の項目群を保持している

37

新作項目の特性値の推定方法

項目特性値既知と未知の項目を組み合わせて出題する

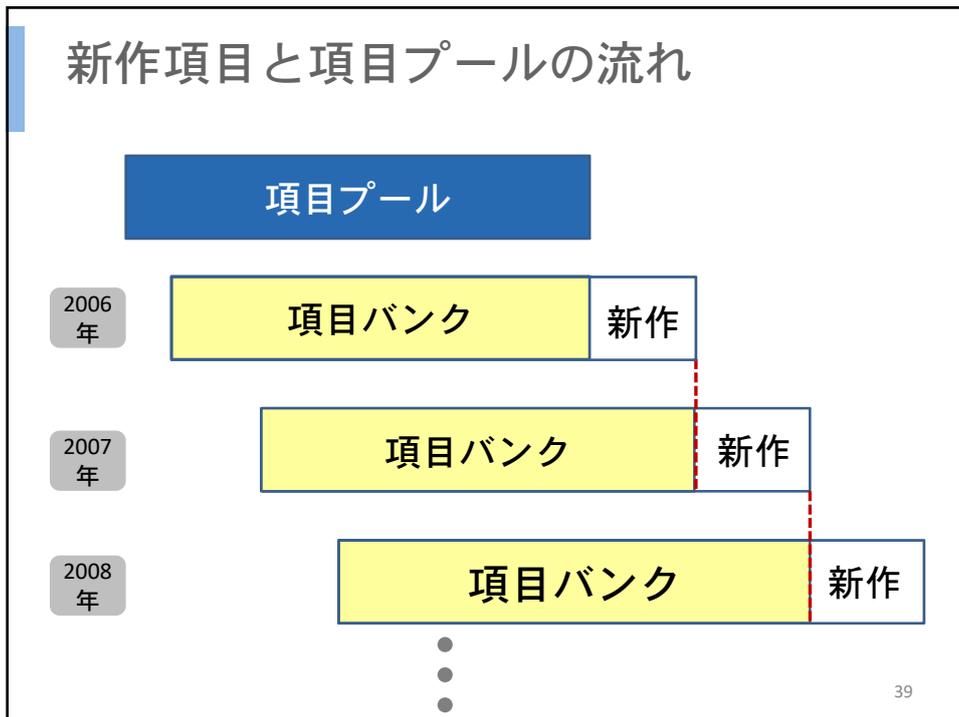


項目特性値既知の項目の解答状況から受験生の能力値を求める



この能力値から項目特性値未知の項目の特性値を推定する

38



基準集団の利用

比較対象とする集団を基準において、経年変化を見ることが可能となる

現在の基準集団

医学系	2012年～2014年の3年間の分布
歯学系	2013年の分布

- 基準集団が平均500、標準偏差100になるように項目特性値を変換する。
- IRT標準スコアは基準集団を通して比較可能となる。

40

これまでの大きな改訂

医学系

- ・ 基準集団を2006年→2012年～2014年

歯学系

- ・ 基準集団：2008年→2013年
- ・ 対象コアカリ：タイプAのみ→全領域

共通

- ・ 項目特性値をコアカリ大分類・問題タイプ別 →全領域
を一つにする
- ・ IRT標準スコアの表示を2桁→3桁

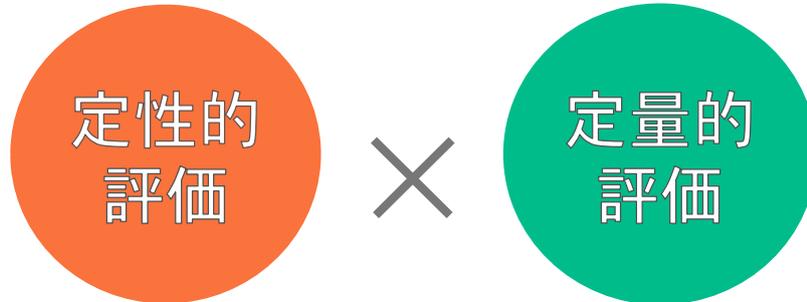
注 継続的な評価のために、等パーセントイル法等の対応

41

項目反応理論を用いた良質な プール問題の選択

42

基本的な項目取捨選択基準



問題の質についての評価
(定量的データを踏まえた)
・問題評価・プール化専門
部会

項目プールからの除外
例: 正答率、双列相関
係数
など

43

項目の取捨選択基準に用いられる統計値

- 項目特性曲線のパターン
- 正答率、双列相関係数
- 項目識別力、項目困難度
- 露出数
- 選択肢別回答率
- 選択肢別解答回

44

IRTを用いた項目取捨選択基準

標準誤差

項目困難度, 項目識別力の信頼性の評価である標準誤差が基準値を超えた場合(推定した項目特性値の信頼性不足)

経年変化

年ごとに新たに求めた項目特性値と固定した項目特性値を比較し、有意差があった場合(急激な変化)

2～3年状況が継続すればプール項目から除外

45

項目反応理論を用いた
共用試験の結果と効果

46

共用試験CBTの項目反応理論を用いた結果について

- CBTにおけるクリティカルポイントの誤差が小さい。
- 平均値が数年にわたり安定的である。
- 問題評価において、直感的に分かる評価となり、安定的な問題評価が可能である。
- 基本的な問題セットに有効利用が可能である。

47

試験信頼性向上専門部会

- テストの専門家（現在、7名）と各委員会の代表を含む委員会
- 各テストの信頼性についてデータに基づく検証を行っている。
- 項目反応理論、一般化可能性理論などの統計理論を用いた手法の正当性の検討
- 各委員会への信頼性に関わる情報のフィードバック
- テスト理論を含む講演会の開催

48

共用試験の効果

1. モデル・コア・カリキュラムの導入、臨床推論に関わる講義の導入による各大学の授業の内容・カリキュラムの変更がみられる
 2. 実習における学生の技能面・態度面（挨拶、問診）の改善が見られるとの報告がある。
 3. 国家試験との相関が高く、早い段階での評価が可能
 4. ある特定の教師の特殊な問題というのがなくなり、きちんとした到達度評価として有用であるとの指摘がある。
- 今後の診療参加型実習の充実度の調査などを含めたアプローチにより、共用試験の妥当性が評価されていくと考える

49

まとめ

50

まとめ

- 共用試験とは全国医学部・歯学部の111学部
の連合体（公益社団法人医療系大学間共用試
験実施評価機構）が運営・実施する。
- 診療参加型実習の充実を大きな目的として、
到達度評価として、有用性が評価されている。
- 試験結果は想定した得点にほぼ到達しており、
導入の成果は上がっている。

51

まとめ

- 技能／態度を評価する試験、臨床推論を検討
する試験を含めることにより、各大学のカリ
キュラム内容にインパクトを与えている。
- コンピュータの特性を利用し、診療の流れを
考慮した新しいタイプの試験が行われ、評価
がなされている。
- さらに公平性があり、信頼性のある試験を目
指して、改善の努力を続けている。

52