

国立大学における教育に関するIRの実践 —千葉大学の事例を中心に—

岡田聡志
千葉大学高等教育研究機構

2017年度第4回千葉大学アカデミック・リンク・セミナー
ALPSセミナー
平成29年度千葉大学全学FD研修会

2018年1月30日（火）14:00-16:00
於：千葉大学ALC

1

本日の構成

- 現在のIRを取り巻く状況の整理
- 目的の再確認
- 全学としてのIRの立ち位置の検討
- どのように具体的な行動に結びつけるか
- まとめ

2

内部質保証と評価の効率化

- 機関別認証評価第3サイクルにおける**内部質保証の機能的有効性**の強調
 - 内部質保証
「PDCAサイクル等の方法を適切に機能させることによって、質の向上を図り、教育・学習その他のサービスが一定水準にあることを大学自らの責任で説明・証明していく学内の恒常的・継続的プロセス」
(大学基準協会『大学評価ハンドブック』)
「大学が自律的な組織として、その使命や目的を実現するために、自らが行う教育及び研究、組織及び運営、ならびに施設及び設備の状況について継続的に点検・評価し、質の保証を行うとともに、絶えず改善・向上に取り組むこと」(大学改革支援・学位授与機構、2017)
 - 「**教育プログラム**」の第一義的重要性と分野別評価の導入
 - 日本学術会議 (2010) 「大学教育の分野別質保証の在り方について」
 - JABEE、法科大学院認証評価、JABPE、JACME、日本看護学教育評価機構 (仮) など
 - 海外での機関別認証評価の負担軽減と軽量化
 - アメリカの連邦教育省によるリスクベースアプローチの奨励とテーマ評価パスウェイ
 - ドイツ、スペイン、オランダ、ベルギーなど

3

日本におけるIRの条件的制約

- 中間組織の脆弱性
 - 中央で一元管理する機関も無く、共通の学問レベルの基準も無い中で、アメリカの高等教育を一つのシステムにするのは、無数のあらゆる種類の団体 (例えば、学問分野別、機関の種類ごと、職業団体、運営者団体など) の、命令ではなく説得のための、議論と話し合いの継続によるものである。(トロウ訳書2000)
 - 「タテの統制が強いため、ヨコの連携が育たない」
 - 「親睦団体的性格が強い」(天野2008)
 - 大学支援組織、中間団体の役割の重要性 (羽田2008、金子2012など)
 - 学士課程答申 (2008) における大学団体の役割への言及
- 専門職と専門職団体の不在
 - 職能開発や同業者規制、自己規律のための倫理綱領策定などの課題 (羽田2008、2012)

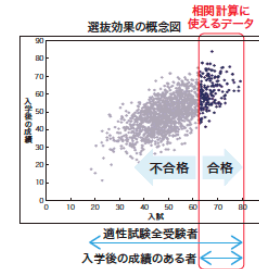
4

事例：入試の予測的妥当性の検証①

- 理由をどう設定するか？
 - 医学教育分野別評価基準 ver.2.2.
医学部は、以下の項目に関して、学生と卒業生の実績を分析するべきである。
背景と状況 (Q7.3.1)、入学時成績 (Q7.3.2)
- 分析枠組みや変数をどう設定するか？
 - 入試に関する変数は、入試区分か、入試得点か、選択科目か。
 - 予測の基準は、初年度GPAか、卒業時の累積GPAか、国家試験得点か、進路選択か。
 - 得点を採用する場合、選抜効果にどう対応するか。
- 解釈をどうするか？
 - 入試と卒業の関連性は、単なるスクリーニング仮説に陥らないか？
 - 相関係数の目安をどう設定するか？

事例：入試の予測的妥当性の検証②

- 選抜効果の問題



選抜効果の例 (中畝：2013)

補正相関係数の例

$$r' = r / \sqrt{s^2 + (1 - s^2)r^2}$$

s^2 : 受験者全員の入試成績の分散に対する合格者全員の入試成績の分散の比

- ※但し、選抜前が正で、選抜後の相関が負である場合、補正公式は存在しない。
 - 特に選抜性 (競争率) の高い学部で、この問題が顕在化しやすい。
 - 相関係数なので、外れ値の影響も受けやすく、数値での議論は難しい
- ex. 0.5だったらどうか、0.6だったらどうか、など。

事例：入試の予測的妥当性の検証③

- 海外ではどのように対応しているのか？
- NACAC (National Association for College Admission Counseling) の調査 (2016)
 - 調査回答機関のうち半数が入試の妥当性の検討を行っており、主にIRが担当している。
 - 機関独自の実施が大半 (約8割) であるが、分析枠組みは多くが、College BoardのACESを参照。
 - ACES (Admitted Class Evaluation Service)
 - アメリカの大学入試で利用される標準テストであるSATなどを開発・実施するCollege Boardが提供する無料の入試データ分析サービス。
 - ACESを通じて、各機関がデータ共有を行うことで、社会保障番号と紐づけられ、その分析結果が約1か月後に提供される。
 - データ共有は、「家族教育権とプライバシー法」(FERPA: Family Educational Rights and Privacy Act) に基づくもので、試験の開発・妥当化・管理運営、学生支援プログラムの管理運営、教育改善などの目的において、同意なしの情報開示が認められている。
 - 基準として採用されているのは「大学の初年度GPA」「SATの得点 (領域別)」「高校内順位」「優等課程やアドバンスト・プレースメント科目の履修数」など。

教育に関するIR担当者の悩みの一例

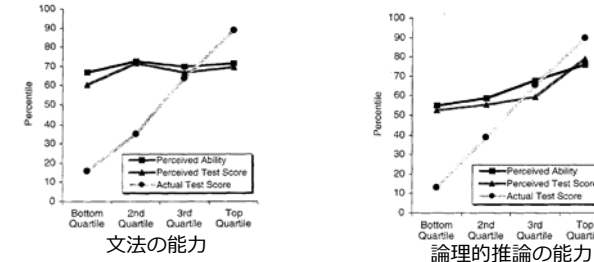


「教育」から分離した「学修成果」の「測定」の難しさ

- 「学修」への特定化の困難。
 - 学生個人の目的意識の違いなど。
- 「学修」によって身に付けた資質・能力はどのように、どの程度維持されるか。
 - 1年次の「学修」は卒業時どうか。
- 自己評価の多義性。
 - 自己評価のメカニズムと解釈の複雑さ
 - ダニング=クルーガー効果
- Gold Standardの不在と相関の解釈の難しさ
 - 進路情報（進学・就職等）は、能力の関係を保証しない。
 - 国家試験は必ずしも資質・能力を保証しない。
 - 標準テストも広範な資質・能力の測定は困難。
- 基本は個別の成績評価

客観評価と自己評価のずれ

- ダニング=クルーガー効果
 - 能力の低い個人ほど、自分の能力を実際よりも高く評価してしまう認知バイアス。優越性の錯覚。



Kruger & Dunning (1999) Figure 3, 4より

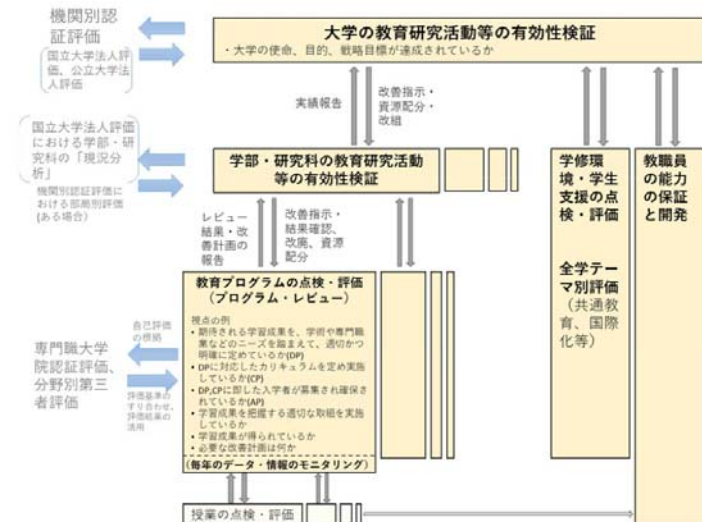
「何のためのIRなのか」の再検討

- 大学における教育の質に対する懐疑
 - 3つのポリシーの設定の義務化（2017年度～）
 - ⇒ 目標設定された資質・能力が、どのような教育・学習によって獲得されるかは、不透明であり、ある種の期待に基づいて行われる。
 - ⇒ 目標中心主義による「目標つぶし」の教育の危険性
 - ⇒ ブラックボックスとして授業（教育目標との対応関係、教育内容、教育方法、教育評価の適切性など）
 - ⇒ “teaching”から“learning”へ



どのようにすれば外側からのチェックは可能か。
教職員や学生個々人の行動が変化しなければ意味がないのではないか。
そこに共通理解はあるか。

内部質保証システムの全体像



大学改革支援・学位授与機構 (2017) 「教育の内部質保証に関するガイドライン」

教育プログラムにフィットするデータの必要性①

- 医学教育を例にすると、
 - 評価基準の粒度
 - 医学教育に関しては、分野別評価基準が細かく、かつ踏み込んだ内容。
 - 専門分野固有の課題やトレンド
 - モデル・コア・カリキュラム、OBE (Outcome-Based Education)、EPAs (Entrustable Professional Activities)、入試におけるMMI (Multiple Mini Interview)、CCにおけるWBA (Workplace-Based Assessment)、など
 - 教育実践との距離
 - 実践上の課題を反映したデータ収集や分析が中心となる。
 - 卒後とのつながり
 - 卒後の研修や生涯学習を含めた医師像への連続性と一貫性という視点の分析。
 - 取り扱うデータの範囲
 - 共用試験やOSCE、Post-CC OSCE、国家試験など
 - IRが取り扱うデータのサイズの小ささ
 - サンプルサイズが小さくなるので、それに対応した分析や解釈の必要性。

13

教育プログラムにフィットするデータの必要性②

- 医学教育分野別評価の国際的基準から見た課題（奈良2017）
 1. アウトカム基盤型教育が十分に浸透していない
 2. 統合型教育（水平、垂直）が実質化されていない
 3. 体系立てられた行動科学教育が不十分である
 4. 診療参加型臨床実習が十分でない
 5. 形成的評価が十分でない
 6. 能動的学修が活発でない
 7. 教育プログラムが定期的に評価を受けていない（PDCAサイクルが実質的に機能していない）

この種の問題は、全学からの視点では捉えにくい。
ボトムアップである必要性。

14

アメリカにおけるIRの設置類型 (Volkwein, 2008)

	相対的に未発達・分散的	相対的に発達・集権的
相対的に規模小	職人的構造 ・ 5000人以下の機関に多い Ex: Oberlin Coll, Norwich U, Kellogg CC	アドホクラシー型 ・ 北東部のIRの30%以上 Ex: Seattle U, Lincoln U, Illinois Coll
相対的に規模大	分散組織型 ・ 研究大学に多い Ex: UC Berkeley, Penn State, Harvard, Yale	専門職的官僚制 ・ 4人以上の専門職 Ex: U Penn, SUNY, Cal Poly

15

全学IRはどのように関わるか

- 千葉大学における全学としての教育に関するIRの活動
 - 「千葉大学における教育IRの基本的考え方」（2015年3月）
 - 定期的実施する活動
 - GPCA、学修時間、アクティブ・ラーニング経験のモニタリング
 - 入試と成績、成績と進路の関係性
 - 当該年度に焦点化して取り組む課題についての活動
 - 満足度に関する分析
 - 奨学金と成績、図書館利用と学業成績の関係に関する分析
 - 就職先の企業属性と学業成績の関係の分析（試行）
 - 各部局からの要望に基づく活動
 - 入試についての選択科目や基準設定に関する分析
 - GPCAのより詳細な分析

16

どのように教育プログラムにコミットするか

- データ収集
 - 教育プログラム独自の関心や視点が含まれるように対応（分岐の設定）。
 - 全体最適の観点から、効率的にデータ収集ができるように。
- 分析結果の報告
 - できるだけ多様なアクターに報告・周知、意見交換。
 - 対執行部
 - 大学運営会議
 - 対教職員
 - 全学FD、部局FD
 - ALPSセミナー、
 - 部局訪問（各部局の教育担当責任者との意見交換）
 - 対学生
 - あかりんアワー
 - 学生との懇談会
 - その他利害関係者
 - 千葉県高等学校長協会連絡会議など

17

GPCAの場合

- GPCA : Grade Point Class Average
 - 成績評価の科目内平均値 ⇒ 成績評価の偏りの目安
- 想定される偏りの理由
 - 受講者数の規模（極端に小さい）
 - 受講者の特性（学習のレディネス、事前の学力等）
 - 受講者間の相互作用
 - 授業形態（講義、演習、実験、実習等）
 - 履修形態（必修・選択）
 - 目標設定の不適切性**
 - 教育内容と評価方法の非対応性**

一律な基準の設定は、個別の教員の評価行動を歪める危険性がある。

教育方法や教育評価の見直しにつながらない。相対評価の導入に流れる？

十分な理由の把握が必要。果たしてそれがどのように可能か。

18

学修時間の場合

- アメリカやイギリスと比べて学習時間が短い
 - NSSE2017：カーネギー分類R1の大学の場合 = 15.6時間/週
 - HEA-HEPI2017：Russell Groupの場合 = 16.1時間/週
→授業時間よりも授業外学習時間の方が長い
- 教育IRとして提示したデータ
 - 学習時間の状況（部局別、学年別）
 - 前年度の学習時間と当該年度の学習時間の関係
 - 予習・復習の指示や宿題のあった授業の割合（アンケートベース）
 - アクティブ・ラーニングと学習時間の関係
- 提案
 - 予習・復習、課題の指示、2) シラバスへの記載の徹底
 - 課題と連動するアクティブ・ラーニングの導入

19

まとめ

- 教育プログラムの改善のためにIRはなにができるか
 - 問題意識の明確化
 - 内的動機づけの重要性
 - 教職員を対象としたアンケートの実施（設計中）
 - 教育プログラムについての意識・関心の醸成
 - カリキュラム改善と個別授業の透明化
 - 教育開発やFD・SDとの連携、連動
 - 具体的な行動を伴う意味のある情報の提供。
 - 開発・推進と評価は分離すべきという意見もある。
 - データ収集・分析・報告の効率化

20

参考文献

- 羽田貴史編. 2008. 高等教育の市場化における大学団体の役割と課題. 科学研究費補助金研究成果報告書.
- 金子元久. 2011. 大学支援組織の可能性と課題. 大学財務経営研究, 8: 3-24.
- Volkwein, J. F. 2008. The foundations and evolution of institutional research. *New Directions for Higher Education*, 2008(141), 5-20.
- NACAC (National Association for College Admission Counseling) . 2016, Use of Predictive Validity Studies to Inform Admission Practice.<https://www.nacacnet.org/globalassets/documents/publications/research/testvalidity.pdf> (2018年1月24日最終閲覧)
- 奈良信雄. 2017. 医学教育分野別評価の意義と展望. 医学教育 = *Medical education (Japan)*, 48(6), 405-410.
- トロウ,マーチンA.著,喜多村和之編訳. 2000. 高度情報化社会の大学: マスからユニバーサルへ. 2000年玉川大学出版部.
- 中畝菜穂子. 2013. 適性試験管理委員会報告書: 適性試験スコアの妥当性に関する分析, JLF NEWS vol.55.