

薬学コアカリからみたデータサイエンス教育の実装

@医療系データサイエンス教育ワークショップ2024

慶應義塾大学薬学部 医薬品開発規制科学講座

漆原 尚巳

今日の内容

- ☑薬学教育、薬剤師教育における統計学・DSニーズ
- ☑薬学教育モデル・コア・カリキュラム
- ☑慶應義塾大学薬学部における医学統計教育
- ☑解析演習の実際

薬学教育における統計学・DSのニーズ

薬剤師

- 薬学部教育
 - 薬学教育モデル・コア・カリキュラム、卒業研究
- 薬剤師国家試験
- 薬剤師生涯教育
～認定・専門資格

製薬産業

- 創薬研究
- 医薬品開発・治験
- 承認後
 - 医薬安全性・薬剤疫学
 - アウトカムリサーチ

医学研究に関する報告論文の正確な理解・伝達
臨床研究のデザイン、実施、解析、報告

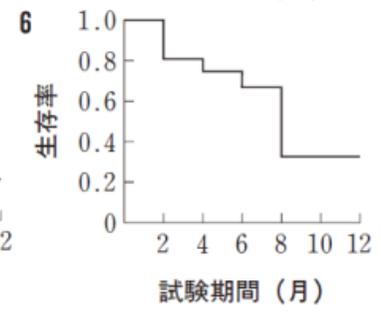
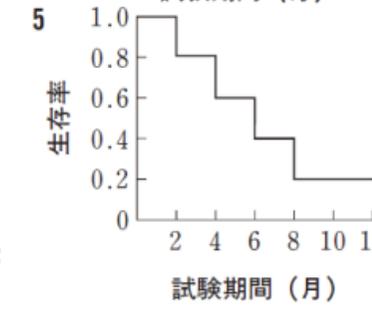
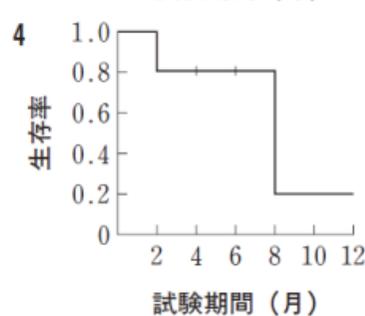
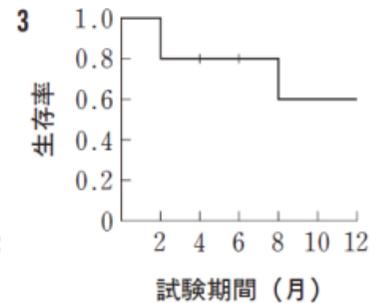
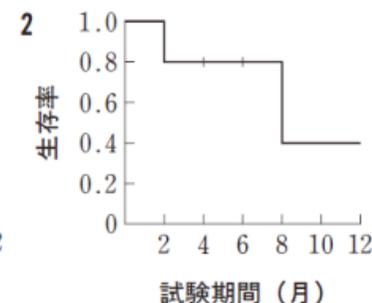
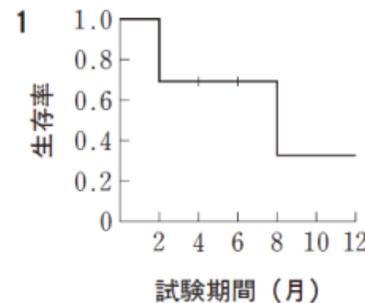
(平成27年実施) 第100回薬剤師国家試験問題

問193

試験期間12ヶ月の臨床試験に参加した5名の被験者の経過が、以下のようになった。

- 1名が2ヶ月後に死亡
- 1名が4ヶ月後に追跡不能となり打ち切り
- 1名が6ヶ月後に追跡不能となり打ち切り
- 1名が8ヶ月後に死亡
- 1名が12ヶ月後の試験終了時まで生存

Kaplan-Meier法を用いて表した生存曲線として、正しいのはどれか。1つ選べ。



論文を必要とする薬剤師の認定・専門資格

日本医療薬学会

- 認定薬剤師、指導薬剤師、薬物療法専門薬剤師、薬物療法指導薬剤師、がん指導薬剤師

日本病院薬剤師会

- 感染制御専門薬剤師、HIV感染症専門薬剤師、精神科専門薬剤師、妊婦・授乳婦専門薬剤師

日本医薬品情報学会

- 医薬品情報専門薬剤師

日本臨床薬理学会

- 認定薬剤師、指導薬剤師

薬学教育モデル・コア・カリキュラム (令和4年度改訂版)

【大項目】

- A) 薬剤師として求められる基本的な資質・能力
- B) 社会と薬学
- C) 基礎薬学
- D) 医療薬学
- E) 衛生薬学
- F) 臨床薬学
- G) 薬学研究

【小項目】

- B-5 情報・科学技術の活用
- D-3 医療における意思決定に必要な医薬品情報
- E-1 健康の維持・増進をはかる公衆衛生

新カリでのDS・統計

- B-5-1 保健医療統計

保健医療統計の種類、データの種類や尺度、データ収集の方法、記述統計及び推測統計

- B-5-2 デジタル技術・データサイエンス

(1)情報科学技術を取り扱う際に必要な倫理観、デジタルリテラシー

(2)医療、保健、介護、福祉におけるデジタル技術・ビッグデータに関連する法規制、デジタル技術の活用例、ビッグデータの活用例

- B-5-3 アウトカムの可視化

(1)薬剤師の活動が社会・地域にもたらす成果(アウトカム)を説明する。

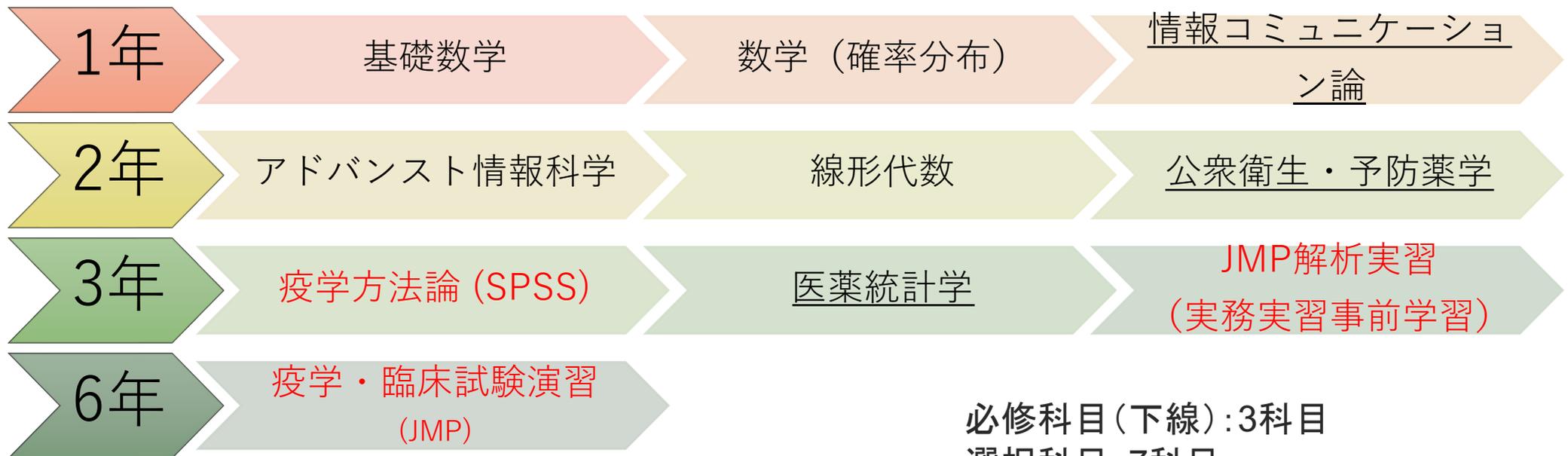
(2)薬剤師の活動の成果を可視化するための測定・評価方法を説明する。

その他DS・統計関連の学修事項

- D-3-3 医薬品情報の解析と評価
 - (2)研究デザインと使用目的、エビデンスの質
 - (4)臨床研究論文の批判的吟味
 - (5)医薬品の有効性評価、安全性評価
- D-3-4 医薬品情報の応用と創生
 - (4)医療ビッグデータの例と特徴
 - (5)不足している情報の創生や課題解決を目指した研究計画
- E-1-1 環境要因によって起こる疾病の予防と健康被害の防止
 - (1)社会や集団における有害事象の発生とリスク因子との因果関係を解析する疫学
 - (4)保健統計及び疫学的手法を用いた疾病や健康被害の背景や原因の解析

慶應薬学部 DS・統計に関する授業

• 旧カリキュラム ~2023



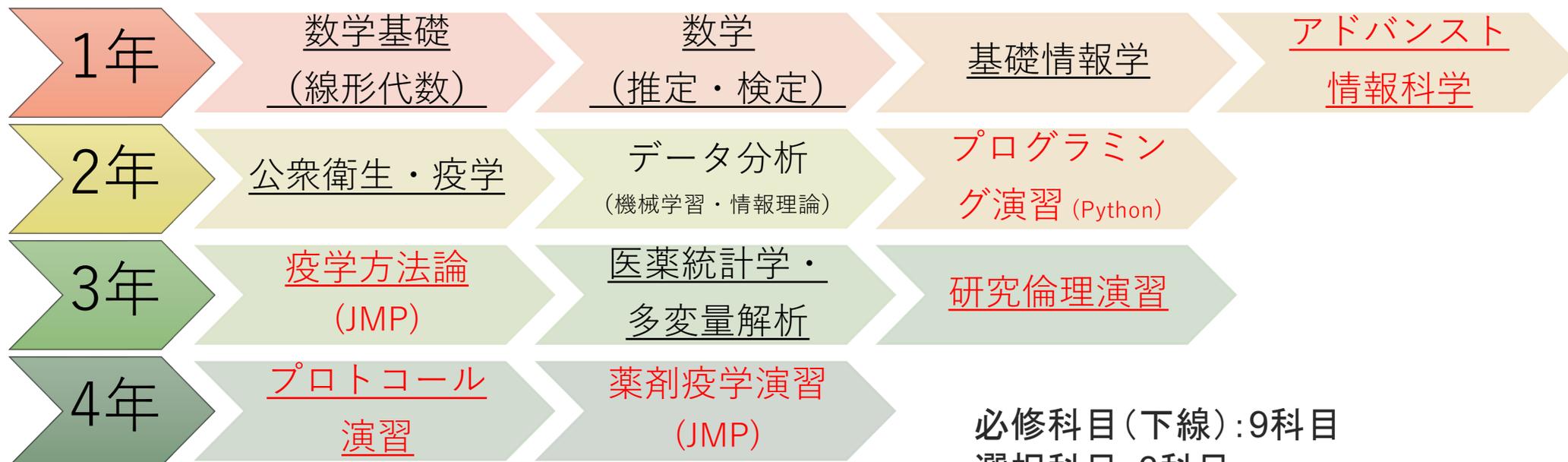
必修科目(下線):3科目

選択科目:7科目

演習科目(赤字):3科目

慶應薬学部 DS・統計に関する授業

• 新カリキュラム 2024~



必修科目(下線): 9科目
選択科目: 3科目
演習科目(赤字): 4科目

医薬統計学 3年生（座学）

回	項目 Topics	到達目標 Attainment Goals	コア・カリSBO
1回	導入（1年次の復習） 母集団と標本	1年次に学んだ基本的な統計量やグラフを説明することが出来る。代表的な離散型・連続型の確率分布を挙げ、母集団と標本の関係を確認率分布の側面から説明することができる。平均値の区間推定を行うことが出来る。	A- (1) -②-4 E3 (1) -⑤-1 E3 (1) -⑤-3
2回	点推定と区間推定	代表的な統計量について、点推定や区間推定を行うことが出来る。2つの母平均の差の区間推定を行うことが出来る。	E3 (1) -⑤-2 ア-E3-③-3
3回	仮説検定	仮説を立てて検定する手法を理解し、1つの母集団について検定することが出来る。	E3 (1) -⑤-2 E3 (1) -⑤-6 ア-E3-③-3
4回	差の仮説検定	2つの母集団について仮説を立て、検定をすることが出来る	E3 (1) -⑤-5
5回	χ^2 検定 分散分析	χ^2 分布を利用した適合度や独立性の検定を行うことが出来る。分散分析の手法を理解し、3群以上の母集団に対して検定を行うことが出来る。	E3 (1) -⑤-5 ア-E3-③-1
6回	ノンパラメトリック検定	パラメトリックとノンパラメトリックの違いを理解し、ノンパラメトリックな場合の検定を行うことが出来る。	E3 (1) -⑤-4
7回	カプランマイヤー法 ログランク検定	カプランマイヤー法を理解し、生存率を求めることが出来る。ログランク検定の手法を理解し、解析することが出来る。	E3 (1) -⑤-7
8回	疫学研究への応用	今まで学んできた統計的手法の現場での応用例を知り、その有効性について理解することが出来る。傾向スコアの医療応用について理解する。	E3 (1) -⑤-1~7

調査方法／解析スキル

- 疫学方法論 現3年生 → 新2年生（必修化）

1	疫学概念、疾病頻度の指標と曝露効果の指標
2	疫学調査法、統計の基礎及びデータ表現、量的データと質的データの収集と分析
3	研究計画・プロトコール
4	調査データの処理・データマネジメント
5	調査法演習 調査票の作成
6, 7, 8	統計解析・演習（代表値やバラツキ、集計方法、相関分析、 χ^2 検定、ノンパラメトリック検定、因子分析）

- JMP解析演習 現3年生 → 新3年生（薬科学科も必修化）

- 実験系の卒業研究にも使えるように配慮

1	リスク比とオッズ比及び95%信頼区間の求め方、 χ^2 検定を理解する。
2	JMPを使い、実際のデータで演習問題を解く。
3	臨床研究に必要な回帰分析を理解する。
4	JMPを使い、実際のデータで演習問題を解く。

解析演習科目の到達目標

1. リスク比とオッズ比の関係を確認する。
2. リスク比・オッズ比の計算、 χ^2 検定を行う。
3. 連続尺度変数の分布を確認し、2群間の比較検定（Student-t検定、Paired t 検定）をする。
4. 回帰分析を理解する。

90分×4コマ 360分

到達目標1. リスク比とオッズ比の関係を確認する。

リスク比を固定したまま、基礎発症率IP0を変化させたときのオッズ比を求める

1. 操作 「ファイル」 「新規作成」 「データテーブル」

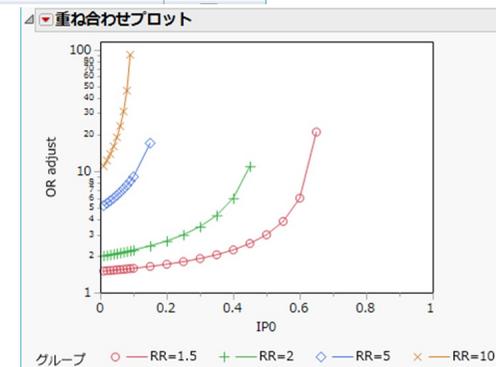
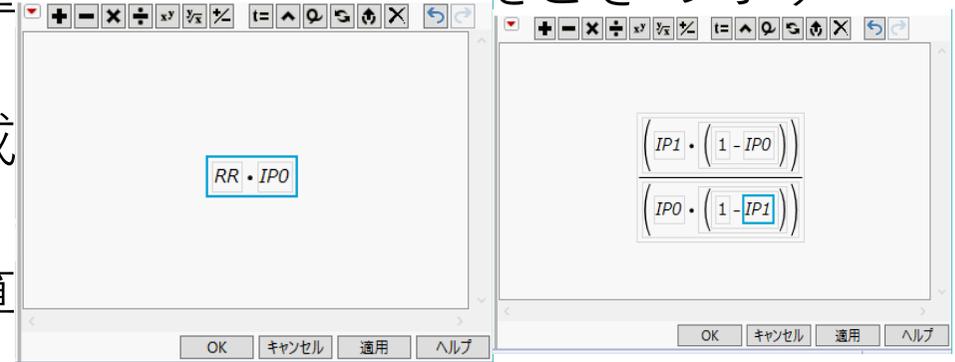
2. 操作 「列の新規作成」 「計算」

- $IP1 = RR \times IP0$

- $OR = \frac{IP1 \times (1 - IP0)}{IP0 \times (1 - IP1)}$

3. 操作 「グラフ」 「重ね合わせプロット」

オッズ比の意味を理解する



到達目標2. リスク比・オッズ比の計算、 χ^2 検定を行う。

	急性膵炎 あり	なし
2型糖尿病患者	117	14,590
非糖尿病患者	356	185,676

リスク比、オッズ比、95%CI、 χ^2 値を手計算で求める

$$\begin{aligned} \text{Log RRの標準誤差} &: \sqrt{\frac{B}{A \times (A+B)} + \frac{D}{C \times (C+D)}} = \sqrt{\frac{14,590}{117 \times (117+14,590)} + \frac{185,676}{356 \times (356+185,676)}} \\ &= \sqrt{\frac{14,590}{1,720,719} + \frac{185,676}{66,227,392}} = \sqrt{0.011282627059378} = 0.1062197112563296 \end{aligned}$$

$$\text{Log RRの95\%Confidential Interval} : \log RR \pm 1.96 \times \sqrt{\frac{B}{A \times (A+B)} + \frac{D}{C \times (C+D)}}$$

$$\chi^2 = \frac{N \times (A \times D - B \times C)^2}{E \times F \times G \times H}$$

重回帰モデルを用いた変数間の調整

- 3年間のフィットネスがHDLコレステロール変化に影響を与える変数を検討する
- 操作 「分析」 「モデルのあてはめ」
 - 多変量回帰モデルによる調整を実施し、解釈する

項	推定値	標準誤差	p値 (Prob> t)	下側95%	上側95%
切片	-1.412	3.226	0.664	-7.937	5.114
性別[female]	3.043	2.454	0.222	-1.921	8.008
運動(分/週)	0.008	0.018	0.660	-0.028	0.044
飲酒	-8.081	3.607	0.031	-15.378	-0.785

予測式： $y = -1.412 + 3.043 \cdot \text{gender} + 0.008 \cdot \text{exercise min} - 8.081 \cdot \text{drink}$

レポート課題

1. (リスク指標) RRとORの関係について考察せよ.
2. どうしたらリスク比が表示されるのでしょうか.
リスク比の計算結果を書きなさい.
また、出てきたリスク比は、何の、何に対する比か、説明しなさい.
3. (2変数の比較) 「対応のない」検定と、「対応のある」検定はどう違うのか、どう使うのか説明せよ.
4. (重回帰分析) HDLコレステロール変化に、有意に影響を与えていた因子は何でしょうか?
 - 計算結果の表を完成させ、予測式を作れ.
 - この結果は、どういう事を意味しているか、説明してください.

医歯薬系応用基礎レベルモデルシラバス 薬学教育モデル・コア・カリキュラム 対比

医歯薬系応用基礎レベルモデルシラバス		薬学教育モデル・コア	慶應薬学シラバス
基礎知識	ベクトルと行列の基礎、記述統計	-	数学基礎（線形代数）、医薬統計学・多変量解析、公衆衛生・疫学
	確率変数と推測統計	B-5-1 保健医療統計	数学（推定・検定）
	回帰分析	E-1-1 環境要因によって起こる疾病の予防と健康被害の防止	数学基礎（線形代数）、数学（推定・検定）、医薬統計学・多変量解析
	統計的検定		数学（推定・検定）、医薬統計学・多変量解析
	探索と知識表現	B-5-3 アウトカムの可視化	公衆衛生・疫学、医薬統計学・多変量解析
	推論と学習	-	データ分析
	教師なし学習の基礎	-	データ分析
	ITセキュリティの原理と法令	B-5-2 デジタル技術・データサイエンス	基礎情報学、アドバンス情報科学
社会動向	データ駆動社会と医療現場	B-5-1 保健医療統計	基礎情報学、公衆衛生・疫学
	医療データの取得と取り扱い	B-5-2 デジタル技術・データサイエンス B-5-3 アウトカムの可視化 D-3-4 医薬品情報の応用と創生	
	AIの歴史と原理	B-5-2 デジタル技術・データサイエンス	データ分析
	統計モデルによる医療データの分析	E-1-1 環境要因によって起こる疾病の予防と健康被害の防止	疫学方法論
	診断におけるAIの活用	B-5-2 デジタル技術・データサイエンス	
	生成AIの動向と仕組み		データ分析
	社会動向についてのグループプレゼンと討論	-	

医歯薬系応用基礎レベルモデルシラバス 薬学教育モデル・コア・カリキュラム 対比

医歯薬系応用基礎レベルモデルシラバス		薬学教育モデル・コア	慶應薬学シラバス
中核項目(1) データサイ エンス	医療で使われるデータサイエンス	B-5-2 デジタル技術・データサイエンス	公衆衛生・疫学、データ分析
	医療で使われるAI		基礎情報学、データ分析
	データ分析の進め方	B-5-2 デジタル技術・データサイエンス	公衆衛生・疫学、データ分析、プロトコール演習 疫学方法論、薬剤疫学演習
	医療データ取り扱いの留意点	B-5-3 アウトカムの可視化	
	医療データの観察	E-1-1 環境要因によって起こる疾病の予防と健康被害の防止	
	医療データ分析の実際	D-3-4 医薬品情報の応用と創生	
	医療データの可視化		
AIの活用と倫理	B-5-2 デジタル技術・データサイエンス	基礎情報学	
中核項目(2) データエン 지니어リン グ	医療におけるICTの進展	B-5-2 デジタル技術・データサイエンス	基礎情報学、アドバンスト情報科学
	医療で使われるAI		
	データ活用の実際	B-5-3 アウトカムの可視化	データ分析
	医療データ収集の実際	B-5-2 デジタル技術・データサイエンス B-5-3 アウトカムの可視化	公衆衛生・疫学、データ分析、プロトコール演習 疫学方法論
	医療データベース作成の手法	E-1-1 環境要因によって起こる疾病の予防と健康被害の防止	疫学方法論
	医療データクレンジングの技術		疫学方法論
	医療倫理とITセキュリティ	B-5-2 デジタル技術・データサイエンス	基礎情報学、研究倫理演習
個別化最適医療を目指したAIの進展と現況	-		

結語

1. 薬学教育における生物統計の重要性
 - 研究のデザイン・データの収集方法（臨床開発・疫学研究）
 - 因果推論（原因究明）に重きを置く
2. ビッグデータ、データサイエンティスト需要
 - 人間、医療制度を出自とするデータの知識・倫理が必要
 - Real World Data（いわゆる2次データ）の薬事利用
3. AI・機械学習の利用
 - 創薬部門，特にインフォマティクスには重要
 - 臨床研究では予測モデルのニーズはそれほど大きくない（因果推論が重要）
 - 現行の医歯薬モデルシラバスで、DS・解析（そのほとんどが予測モデル）に焦点を置いているのは、薬学モデコアに合っていない
4. 薬学（おそらく医学歯学も）は必修科目が多すぎて、これ以上増やせない

