



「AI 解析に基づく十二指腸乳頭分類とERCP におけるアウトカムとの相関性に関する検証」 研究成果報告

2026年7月16日 日本消化器内視鏡学会

研究責任医師:松田 浩二

(国立病院機構 静岡医療センター)

研究協力：九州大学 原田 翔太 助教・備瀬 竜馬 教授・内田 誠一 理事



1. 研究の内容

1) 研究の目的

十二指腸乳頭部の内視鏡画像におけるAI解析に基づく新たな十二指腸乳頭の分類を考案し、医療メタデータ（患者の属性、診断結果、問診情報等）を使用して、ERCPにおけるアウトカムとの相関性を検証し、新分類に基づいた十二指腸乳頭部の内視鏡画像からアウトカムを予測する自動診断の可能性について検証する。

2) 主要評価項目

AI による十二指腸乳頭部の画像分類

3) 副次的評価項目

JED データとのアウトカム(胆管挿管率・偶発症など)における相関関係
十二指腸乳頭部の内視鏡画像データベースの構築

4) 探索的評価項目

乳頭部撮影時の AI による臨床的アウトカム予測



5) 調査項目

調査対象期間: 2015年1月から2019年12月31日までの情報を調査対象とする

患者背景	年齢、性別、ASA Grade、生活歴（喫煙・飲酒）、家族歴、他臓器癌既往歴、抗血栓薬の服用状況
ERCP 関連項目	検査目的、全施行時間、乳頭部までの到達時間、胆道造影範囲、胆管挿管、最初の胆管へのアプローチ法、最終的な深部挿入時のアプローチ法、深部挿入までの乳頭へのアプローチ回数、胆管径、膵管挿管、膵管径、ERCP 難易度（Schutz 分類）、全照射時間、専攻医が施行した場合の完遂率、施行前の治療の有無、到達部位、手技前偶発症、手技中偶発症、翌日の amylase 値、手技後偶発症、30 日以内の死亡
内視鏡画像	ERCP 実施医が検査報告書に添付した画像から、胆管挿管を試みる前の十二指腸乳頭の開口部と口側隆起が含まれているより鮮明な画像を抽出する



6) 研究実施機関

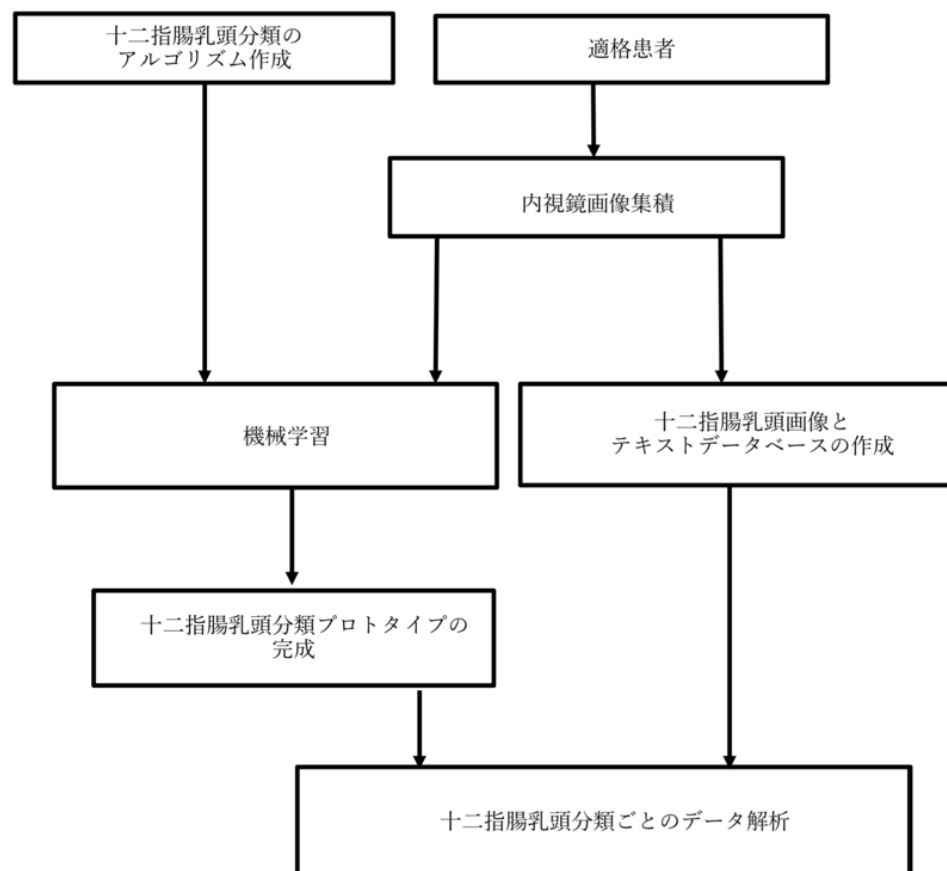
研究実施許可日(通知書発行日)～2026 年 3 月 31 日まで

7) 研究参加施設

施設名	研究責任者	所属・職位
手稲溪仁会病院	潟沼 朗生	消化器内科・消化器病センター長
富山大学付属病院	安田 一朗	第三内科診療部門 消化器内科・教授
独協医科大学病院	入澤 篤志	消化器内科・教授
埼玉医科大学国際医療センター	良沢 昭銘	消化器内科・教授
東京医科大学病院	糸井 隆夫	消化器内科・教授（副院長・診療科長）
東京大学医学部付属病院	中井 陽介	光学医療診療部・准教授
	高原 楠晃	消化器内科・特任講師
東京慈恵会医科大学葛飾医療センター	加藤 正之	内視鏡医学講座・診療副部長
北里大学病院	木田 光広	消化器内科・教授
	奥脇 興介	消化器内科・講師
静岡医療センター	松田 浩二	消化器内科・診療部長（内視鏡センター長）
京都第二赤十字病院	田中 聖人	消化器内科・副部長（医療情報室長）
	鈴木 安曇	消化器内科・医長
	宇野 耕治	消化器内科・部長
久留米大学病院	岡部 義信	消化器内科・准教授
	寺部 寛哉	消化器内科・助教

※研究責任者変更時点（変更がない場合は研究開始当初）の所属・職位となっております。

8) 研究の流れ





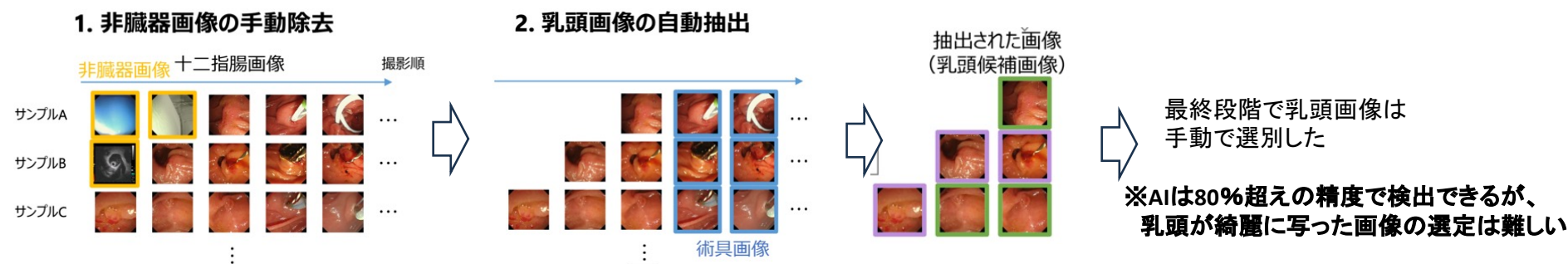
2. 研究の成果報告

1) ERCP検査画像の収集結果

十二指腸乳頭AI研究 ERCP検査画像収集実績		2022.8現在
ご施設名	検査数	画像数
埼玉医科大学国際医療センター	2,523	54,508
東京大学医学部附属病院	662	7,339
国立大学法人富山大学附属病院	1,002	25,810
医療法人涇仁会手稻涇仁会病院	6,953	112,247
獨協医科大学病院	4,982	82,199
学校法人慈恵大学東京慈恵会医科大学葛飾医療センター	1,538	32,403
北里大学病院	4,034	91,696
久留米大学病院	1,648	65,190
学校法人東京医科大学東京医科大学病院	5,049	62,374
日本赤十字社京都第二赤十字病院	809	17,369
合計	29,200	551,135

2) AIプロトタイプにトライアル結果

①【乳頭画像の自動抽出】 処置具が挿管される前の画像を乳頭候補画像としてAIが抽出

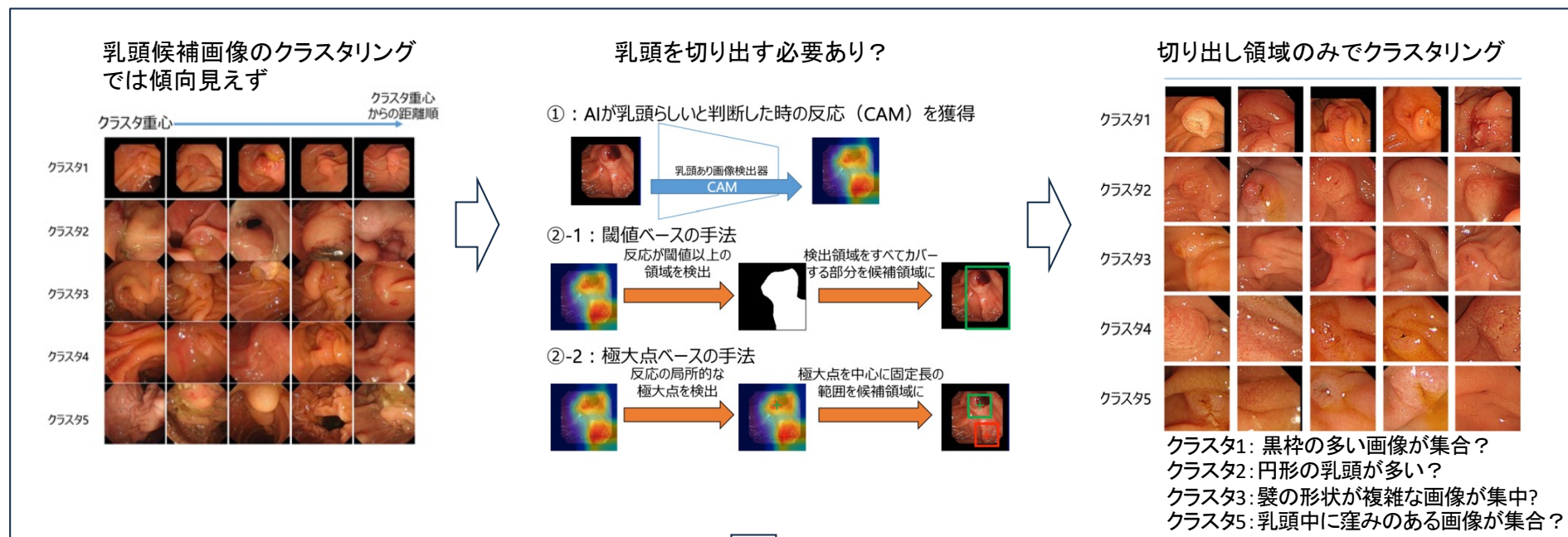


②【乳頭画像からの難易度判定】 手動で選定した画像に挿管時間を付与し、画像から難易度予測を試験



※ 図表は、九州大学原田 翔太 助教・備瀬 竜馬 教授・内田 誠一 理事の発表資料から引用

③ 【乳頭タイプ別クラスタリング】



※乳頭形状(別開口型、タマネギ型、結節型、絨毛型、平坦型、縦長型)の分類と乖離がある

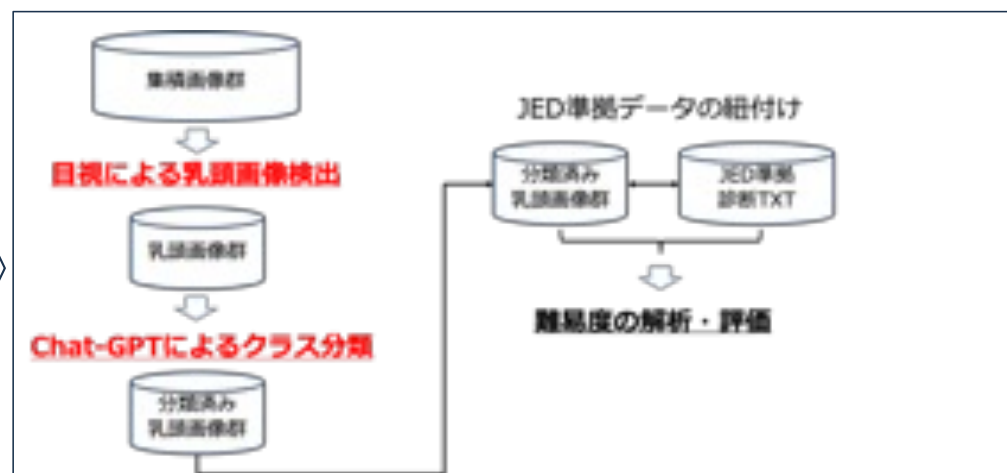
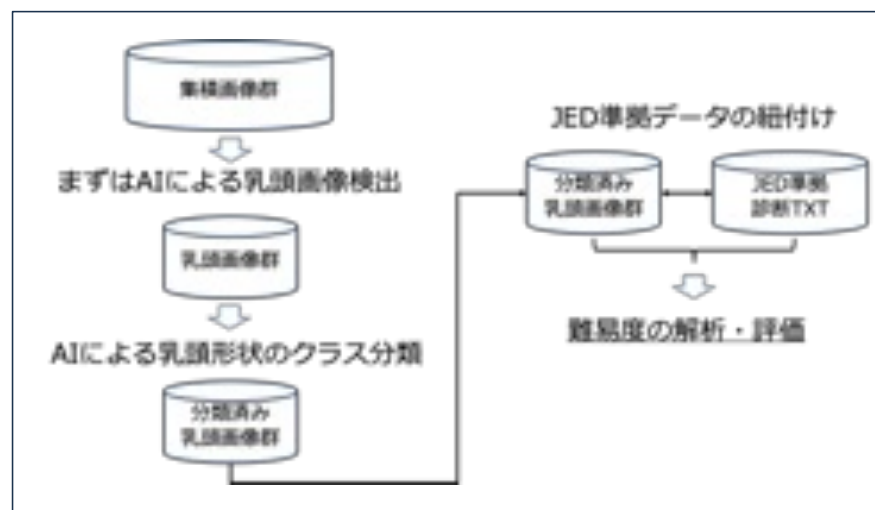
※ 図表は、九州大学原田 翔太 助教・備瀬 竜馬 教授・内田 誠一 理事の発表資料から引用

7

First of all, For patient safety, For all of endoscopist. Concept by JGES

3) アルゴリズムの変更による課題解決

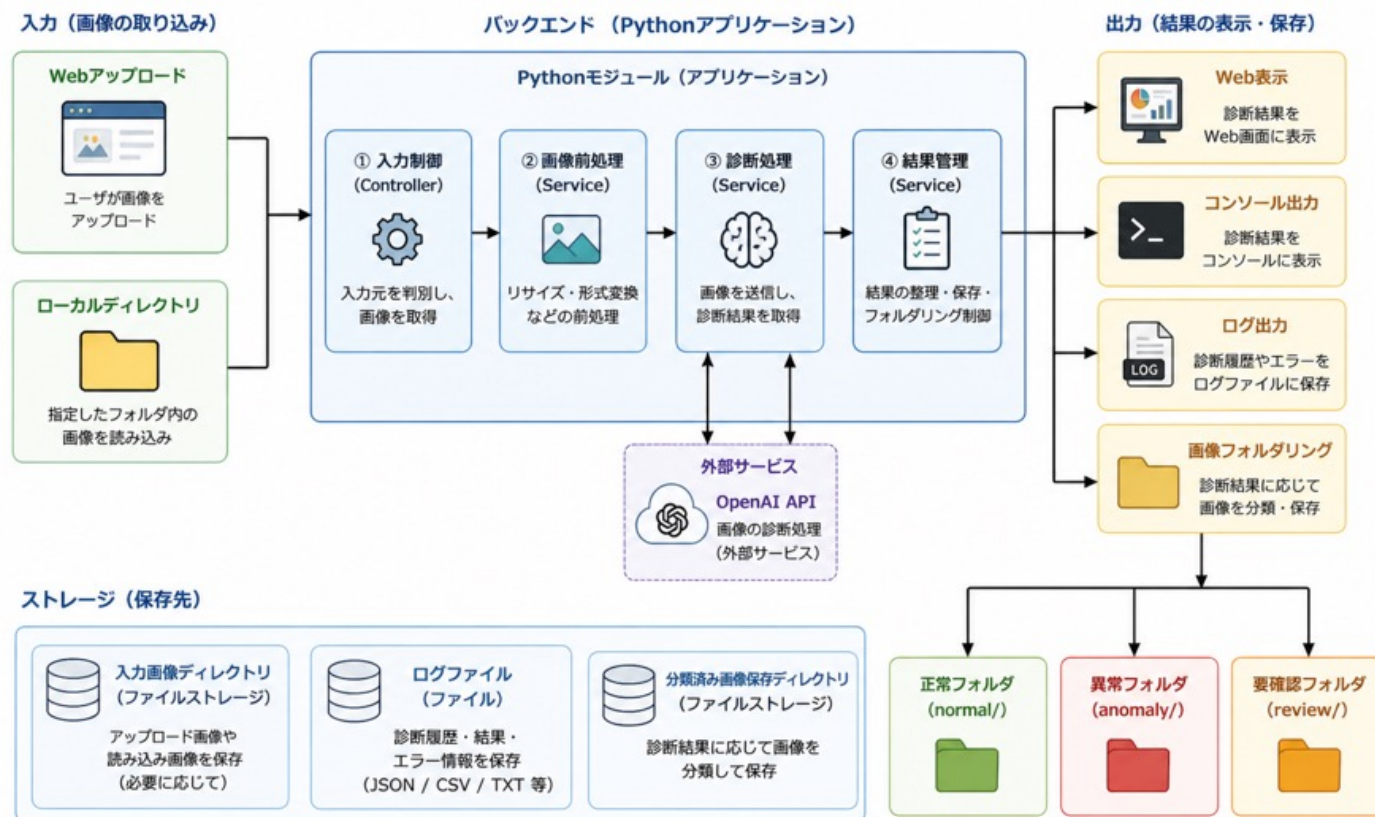
九州大学（原田 翔太 助教・備瀬 竜馬 教授・内田 誠一 理事）との協議により、近年著しく医療領域でも性能が向上しているオープンAI Chat-GPTを利用する方針に変更



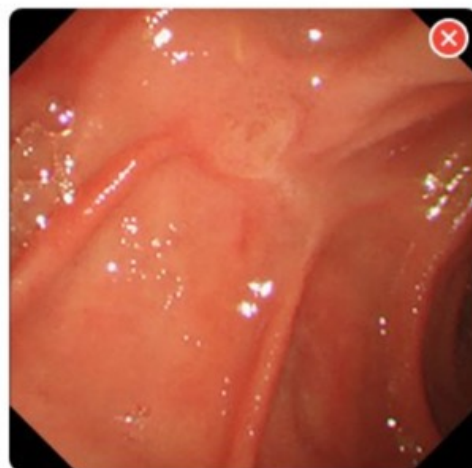
新たにクリアすべき技術課題

- ・セキュリティ対策 — Chat-GPT チューニング 本件以外に学習に使用しない設定する
- ・大量のデータ処理 — 通常対話型利用、大量のデータを処理すべくAPIを開発する

4) 開発したオープンAI-APIとアプリケーション全体図

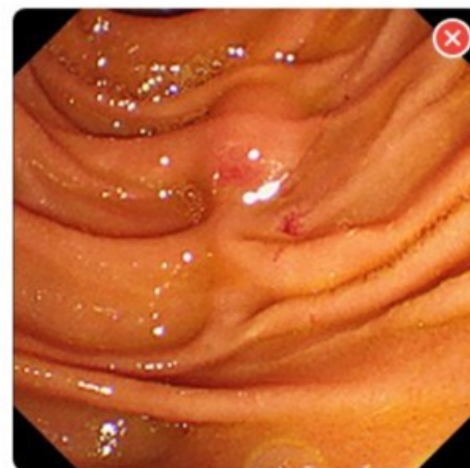


5) 乳頭タイプ判定処理結果サンプル



分類: Type 1 (平坦型)

根拠: 表面が平坦で突出がないため。



分類: Type 3 (乳頭型)

根拠: 明確な乳頭状の突出が見られるため。

Chat-GPT回答
画像を送り、判定
タイプと判定理由
を返す設定



6) JED準拠データの紐付け方法

1. 対象検査の抽出
 - ・対象施設のうち、テキストの検査IDと画像の検査ID部分を比較する
 - ・検査IDがマッチする検査の中で、ERCP基本情報の乳頭へのアプローチ回数の記載あり検査を対象とする
2. 検査難易度を追加
 - ・対象検査に対し、基本情報を元に、テキストへ難易度を追加
3. 乳頭画像のピックアップ
 - ・対象検査の画像のうち、乳頭に対し何かしら挿入される1つ手前の画像を対象とする
 - ※ただし、検査単位の画像のうち、1枚目から何かしら挿入されている場合は、検査自体を対象外とする。
4. テキスト難易度と画像Typeを比較
 - ・画像TypeのAI判定結果が、1回目と2回目で異なる画像は対象外とする
 - 例) 同じ画像に対し2回チェックを行い、
 - 1回目はType1、2回目もType1 → 対象
 - 1回目はType1、2回目はType2 → 対象外



7) 画像とJED準拠データの紐付け結果

テキストと画像の検査ID突合チェック

施設名	テキスト	画像		テキストと画像の検査ID比較	
	検査数	検査数	枚数	突合結果	不一致理由
手稲溪仁会病院	6,848	5,447	131,444	一致	
富山大学	1,002	999	29,774	不一致	テキスト側に検査IDがない
北里大学	4,033	3,222	95,655	不一致	検査IDと一致するHASH値作成できず
埼玉医科大学国際医療センター	2,521	2,526	69,638	不一致	検査IDと一致するHASH値作成できず
獨協医科大学	4,982	4,982	101,118	不一致	検査IDと一致するHASH値作成できず
東京医科大学	4,976	4,988	62,374	一致	
慈恵葛飾医療センター	1,538	1,538	41,827	一致	
久留米大学	1,648	1,647	64,612	一致	
東京大学	1,148	662	7,339	不一致	テキスト側に検査IDがない
京都第二赤十字病院	532	532	10,953	一致	
合計	29,228	26,543	614,734		

乳頭へのアプローチ回数記載ありの検査画像抽出

施設名	テキスト		画像
	対象検査数	画像あり検査数	有効画像あり
手稲溪仁会病院	1,489	1,380	729
東京医科大学	418	417	169
慈恵葛飾医療センター	971	971	429
久留米大学	2	—	—
京都第二赤十字病院	532	532	415
合計	3,412	3,300	1,742



8) 新アルゴリズムによる解析結果

施設名	画像フォルダ数 合計	チェック後検査 数	対象画像枚数	テキスト一致画像
	(対象検査数)			
手稲溪仁会病院	1,380	729	729	729
慈恵葛飾医療センター	971	429	429	429
東京医科大学	417	169	169	169
京都第二赤十字病院	532	415	415	415

2回目 1回目	Type 1 (平坦型)	Type 2 (顆粒型)	Type 3 (乳頭型)	Type 4 (腫瘍型)	Type 5 (陥凹型)	合計	Type一致	容易	やや難	難	有効難易度 のみ合計	容易	難	やや難	難
Type 1 (平坦型)	60	42	45	1	4	152	60	25	2	25	52	48.1%	51.9%	3.8%	48.1%
Type 2 (顆粒型)	12	110	94	9	6	231	110	45	10	22	77	58.4%	41.6%	13.0%	28.6%
Type 3 (乳頭型)	45	92	949	47	25	1,158	949	424	83	263	770	55.1%	44.9%	10.8%	34.2%
Type 4 (腫瘍型)	1	9	60	33	5	108	33	10	4	9	23	43.5%	56.5%	17.4%	39.1%
Type 5 (陥凹型)	6	8	28	8	43	93	43	25	3	8	36	69.4%	30.6%	8.3%	22.2%
	124	261	1,176	98	83	1,742	1,195	529	102	327	958	55.2%	44.8%	10.6%	34.1%



十二指腸乳頭のTypeと検査難度の関係性解析

原田 翔太(九州大学)



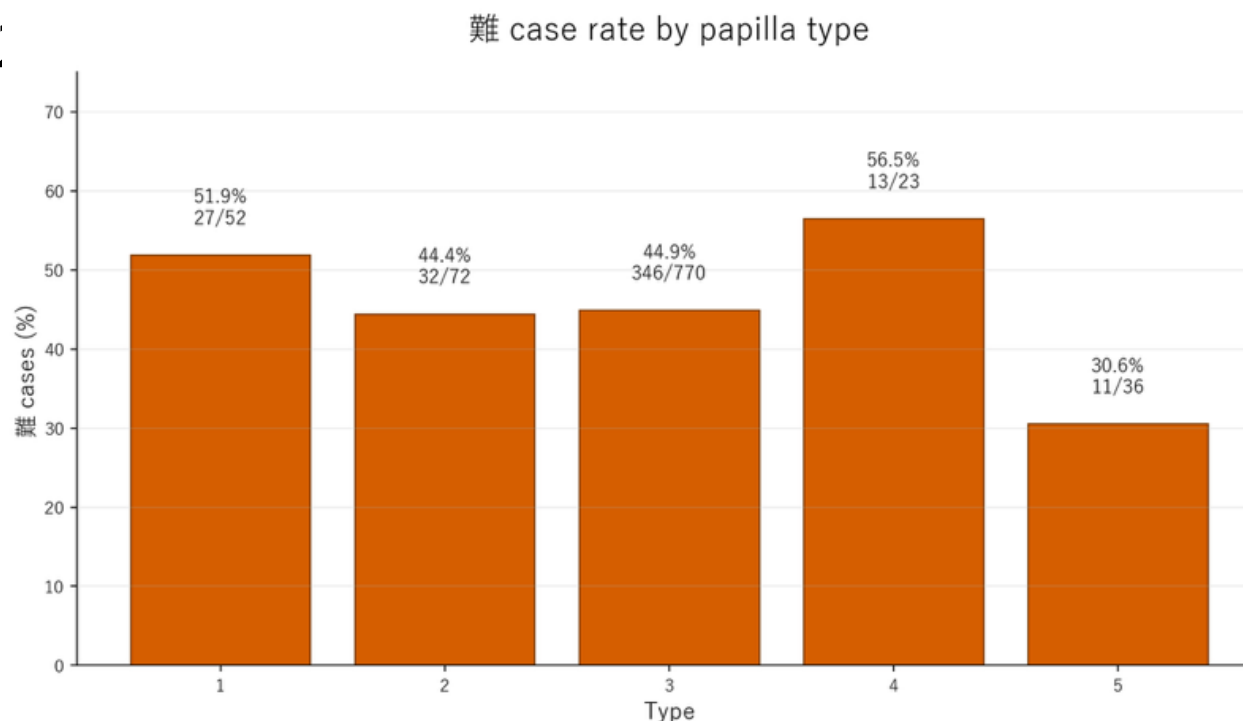
なにをしたのか：難症例率にずれがあるかを検定

- 全体比較 (p. 3)
 - 5つのType (Type 1: 平坦, Type 2: 顆粒型, Type 3: 乳頭型, Type 4: 腫瘍型, Type 5: 陥凹型) 間で難症例率に差があるかを検定
 - 検定方法: カイ二乗検定, 独立性の検定
 - 補正方法: なし
- Type別の比較 (p. 4)
 - 各Typeそれぞれの難症例率が全体の難症例率からずれているか検定
 - 検定方法: カイ二乗検定, 適合度検定
 - 補正方法: Holm補正
- ペア比較 (p. 5)
 - 2つのType間で難症例率に差があるかを検定
 - 検定方法: フィッシャーの正確確率検定
 - 補正方法: Holm補正
- 結論: どの検定でも有意な差は確認されなかった



全体比較：5つのType間で難症例率に差があるかを検定

- 全体比較では有意な差は確認されなかった
 - カイ二乗統計量=5.285
 - p値=0.1





Type別の比較：各Typeそれぞれの難症例率が全体の難症例率からずれているか検定

- 全体の傾向から有意なずれを示すTypeは無かった

各Typeにおける統計量

Type	難症例数/全症例数	期待難症例数	p値 (Holm補正)
Type 1	27/52	23.4	1.00
Type 2	32/72	32.4	1.00
Type 3	346/770	346.6	1.00
Type 4	13/23	10.4	1.00
Type 5	11/36	16.2	0.41



ペア比較：2つのType間で難症例率に差があるかを検定

- 有意な差のあるペアは確認されなかった

p値の小さい3ペアにおける統計量

ペア	難症例率 [%]	オッズ比	p値 (Holm補正)
Type 1 vs Type 5	51.9 vs 30.6	0.407	0.526
Type 4 vs Type 5	56.5 vs 30.6	0.338	0.547
Type 3 vs Type 5	44.9 vs 30.6	0.539	0.971