



Title	膵頭十二指腸切除術のトレーニングプログラムの開発に向けた膵空腸吻合の技能評価システムの構築
Author(s)	水沼, 謙一
Citation	北海道大学. 博士(医学) 甲第15652号
Issue Date	2023-09-25
DOI	10.14943/doctoral.k15652
Doc URL	http://hdl.handle.net/2115/90978
Type	theses (doctoral)
Note	配架番号 :
File Information	MIZUNUMA_Kenichi.pdf



[Instructions for use](#)

学位論文

膵頭十二指腸切除術のトレーニングプログラムの
開発へ向けた膵空腸吻合の技能評価システムの構築

(Development of skill evaluation system for
pancreaticojejunostomy toward the training
program for pancreatoduodenectomy)

2023年9月

北海道大学

水沼 謙一

学位論文

膵頭十二指腸切除術のトレーニングプログラムの
開発へ向けた膵空腸吻合の技能評価システムの構築

(Development of skill evaluation system for
pancreaticojejunostomy toward the training
program for pancreatoduodenectomy)

2023 年 9 月

北海道大学

水沼 謙一

目次

発表論文目録および学会発表目録	1 頁
要旨	2 頁
略語表	5 頁
緒言	6 頁
方法	15 頁
結果	26 頁
考察	41 頁
総括および結論	45 頁
謝辞	46 頁
利益相反	47 頁
引用文献	48 頁

発表論文目録および学会発表目録

本研究の一部は以下の論文に発表した。

1. Mizunuma K, Kurashima Y, Poudel S, Watanabe Y, Noji T, Nakamura T, Okamura K, Shichinohe T, Hirano S. Surgical skills assessment of pancreaticojejunostomy using a simulator may predict patient outcomes: A multicenter prospective observational study. *Surgery*. Jun; 173:1374-1380. (2023)

本研究の一部は以下の学会に発表した。

1. Kenichi Mizunuma, Yo Kurashima, Saseem Poudel, Kotoe Kiriyama, Tomoko Mizota, Yusuke Watanabe, Takehiro Noji, Toru Nakamura, Keisuke Okamura, Toshiaki Shichinohe, Satoshi Hirano.

Development of the assessment and training tools for the pancreatojejunostomy

The Association for Surgical Education (ASE) 2021 Annual meeting, April 2021. Virtual

2. 水沼 謙一、倉島 庸、サシーム パウデル、渡邊祐介、田中公貴、中西喜嗣、浅野賢道、野路武寛、海老原裕磨、村上壮一、中村 透、土川貴裕、七戸俊明、平野 聡

膺空腸吻合における技能の評価ツールおよびシミュレーションモデルの開発

第120回北海道臨床外科学会. 2021年9月・札幌

要旨

【背景と目的】

膵頭十二指腸切除術（Pancreatoduodenectomy: 以下 PD）は、膵頭部周囲の腫瘍を切除するため胃十二指腸、膵臓、胆管の一部を同時に切除し、複雑な再建を要する。PD の合併症発生率は 35～51% と他の術式と比較し高率に発症し、特に膵液瘻は 13～20% の発生率である。膵液瘻の治療が奏功せぬ場合、仮性動脈瘤からの出血が致命的となる恐れがある。致命的な合併症を防ぐため膵液瘻の低減が必要である。

膵液瘻の低減のために、膵液瘻リスク評価や再建方法の工夫が行われてきた。近年では、手術中の膵空腸吻合を技能評価し、評価スコアが低い執刀群は術後の膵液瘻発症が優位に高いという結果が報告され、外科医の技能は患者の予後へ寄与する可能性が示唆されてきた。

外科医の技能向上のためシミュレーショントレーニングの有効性が報告されるなか、PD のシミュレーターも開発されてきたが、臨床上の成績は評価していない。

PD は高度な技術を要する術式であり、技術習得には他の術式と比べ時間を要する。本邦において、PD の標準的な指導方法やトレーニングカリキュラム、シミュレーショントレーニング法は確立されていない。そこで、PD の技能評価スケールおよびシミュレーショントレーニングプログラムによる効率的な修練が、術後の成績を改善するのではないかと考えた。

本研究では、PD のトレーニングプログラムの作成へ向け、術後の合併症に大きく影響する膵空腸吻合をトレーニングするための技能評価スケールを開発し、教育的有効性を検討することを目的とした。

【対象と方法】

既存の評価スケールとして、手術中の一般手技を評価する **Objective structured assessment of technical skill**（以下 OSATS）、ロボット支援下手術の PD における膵空腸吻合の技能評価スケールが存在する。これらのスケールを参照し、肝胆膵外科高度技能専門医および外科教育専門家の意見を統一し、手術室、シミュレーション環境下どちらにおいても評価可能な膵空腸吻合技能評価スケール **Pancreaticojejunostomy performance assessment scale**（以下 PJPAS）を開発した。

本研究の評価方法は研究対象の外科医が PD の執刀を行い、手術室でビデオ撮影した。執刀の 1 か月以内に膵空腸吻合のシミュレーター手技をビデオ撮影した。肝胆膵外科高度技能医 3 名が評価者となり、手術室環境のビデオ（**Operating room video**: 以下 ORV）とシミュレーション環境下のビデオ（**Simulator video**: 以下 SimV）を PJPAS と OSATS の評価ツールを用いて評価した。

シミュレーターは国内で販売されている膵臓・空腸モデルから Biotexture™, (FASOTEC, Japan)を採用した。手術室環境を再現するために、シミュレーターの設置、ビデオカメラの撮影場所の標準位置を設定した。シミュレーターのビデオ撮影は研究者がすべての手技に立ち合い撮影条件を確認した。対象の外科医にシミュレーターについての再現性など評価のアンケートを実施した。

評価者間信頼性を担保するために外科教育専門家2名と評価者となる肝胆膵外科高度技能専門医3名で2時間の評価者トレーニングを行った。評価表の項目および評価基準を詳細に確認した。

本研究の対象者は北海道内の肝胆膵外科修練施設もしくは、年間PD数が10例以上見込める施設に所属する外科医を対象とした。

主要評価項目は、手術室とシミュレーション環境下での膵空腸吻合技能評価スコアとの相関である。副次的評価項目は、評価スコアと術後成績の相関、および評価スコアと外科医の経験との相関である。術後成績は術後在院日数、術後合併症を記録した。術後膵液瘻 (Postoperative pancreatic fistula 以下 POPF) の有無および重篤な合併症の有無も記録した。

【結果】

PJPAS を開発し、評価のばらつきをなくすため評価基準表を別に作成した。研究参加者のアンケートからシミュレーターが実際の手術を高く再現していると評価された。

ORV の3名の評価者の級内相関係数は PJPAS 評価では評価者間のどの組み合わせでも信頼性が低く、OSATS 評価では概ね高い信頼性を認めた。ORV を PJPAS で評価することは信頼性が担保されないとした。SimV における3名の級内相関係数は PJPAS、OSATS どちらの評価スケールでも評価者間の信頼性は担保されていた。

手術室とシミュレーション環境下の OSATS によるビデオ評価では両環境下の技能評価スコアに相関性は認めなかった。

北海道内11施設から30名の外科医が本研究に参加し、ビデオデータ不備等により3名の外科医がデータ解析対象から除外された。27名をPDの執刀経験数に応じて Novice 群 (PD 執刀経験が10例以下)、Intermediate 群 (11例以上、50例以下)、Experienced 群 (50例以上) の3群に分けた。それぞれの群における患者背景に差は認めなかった。3群間において OSATS 評価スコアでは3群間に有意差を認めたが、術後成績では差を認めなかった。

膵液瘻の発症について術後結果と評価スコアを検討因子として単変量、多変量解析を行ったところ、シミュレーターにおける PJPAS スコアの低値が膵液瘻発症因子となりうる結果であった。シミュレーション環境下において OSATS 及び PJPAS 評価で PD の執刀経験数と中等度の相関を認めた。

【考察】

手術室とシミュレーション環境下における評価結果が相関性を認めなかった主要因は、手術室のパフォーマンス評価の難しさにある。膜空腸吻合のような高度な技能を手術室環境で評価する際に信頼性の確保が困難であることや、手術中に使用する器具や環境の違い、患者の疾患や体格、膜臓の状態などの背景の違いが交絡因子となることが考えられた。

シミュレーション環境は一定の条件で手技を行うことができることから、評価の信頼性が確保されていたと考えられる。患者の臨床転帰の分析では、シミュレーション環境下の PJPAS スコアは患者の POPF 発生率と相関を示した。シミュレーション環境下における OSATS、PJPAS 評価は PD の執刀経験数と中等度の相関を認めたことから、PD 習熟度の指標となる可能性がある。シミュレーション環境を利用して術者の技能を手術前にある程度予測することが可能で、手術前の到達目標の基準やトレーニングとして、有効活用できると考えられる。

【結論】

膜空腸吻合技能を手術室、シミュレーション両方の環境下で評価可能なスケールを開発した。膜空腸吻合のシミュレーター環境を設定し、研究参加者よりトレーニングへ有効活用を期待する評価が得られた。高難度手技である膜空腸吻合の手術室における評価法にはさらなる検討の必要があるものの、事前にシミュレーション環境下で術者の技能を測定することができれば、対象術式への習熟度の予測が可能となり、トレーニングで補填できる可能性がある。また、事前の評価スコアは達成目標のマイルストーンとなり得るだけでなく患者のアウトカムを予測できる可能性がある。

略語表

本文および図中で使用した略語は以下の通りである。

BMI	Body Mass Index
FRS	Fistula Risk Score
ICC	Intraclass correlation coefficients
ISGPF	International Study Group of Pancreas Fistula
OR	Operating room
ORV	Operating room video
OSATS	Objective structured assessment of technical skill
PD	Pancreatoduodenectomy
PJ	Pancreaticojejunostomy
PJPAS	Pancreaticojejunostomy performance assessment scale
POPF	Postoperative pancreatic fistula
Sim	Simulation
SimV	Simulator video

緒言

膵頭十二指腸切除術 (Pancreatoduodenectomy: 以下 PD) は、膵頭部周囲の腫瘍を切除するための術式であり、20 世紀初頭に世界で初めて実施され、その後 Allen Whipple が複数の報告例を発表し術式の改良が進められてきた (Whipple et al., 1935)。この術式は、膵頭部周囲の解剖学的特性により十二指腸乳頭部、下部胆管に発生した腫瘍を切除する際にも同様の術式が選択される。胃から十二指腸、膵臓、胆管の一部を同時に切除したのちに、複数の再建を行う。その再建は残胃からの消化管の経路、胆管の断端からの胆汁の経路、膵臓の断端からの膵液の経路を、挙上した小腸とそれぞれ吻合を行う (図 1)。PD は腫瘍の切除から再建に至るまで多くの手順を要し、消化器外科領域では高難度に分類される術式である。

このような肝胆膵外科領域の高難度手術を安全に行うことができるよう、日本肝胆膵外科学会は 2008 年より高度技能専門医制度を発足した。学会が認定した修練施設と指導医のもとで教育を受け、手術実績や手術ビデオ評価を経て高度技能専門医が認定される。修練施設は高難度に分類される肝胆膵領域の手術の症例数と、指導医が常勤していることが条件とされる。専門医の申請には認定基準に定められた手術実績やセミナーの受講が必要である。実臨床の手術中の教育は各施設の方針に委ねられているのが現状である。

同学会のデータによると、本邦の PD は年間約 8000 例行われており、その約半数が修練施設で行われている。修練施設で行われた PD の術後成績は、修練施設以外と比較し良好であると報告されている (Miura et al., 2016)。

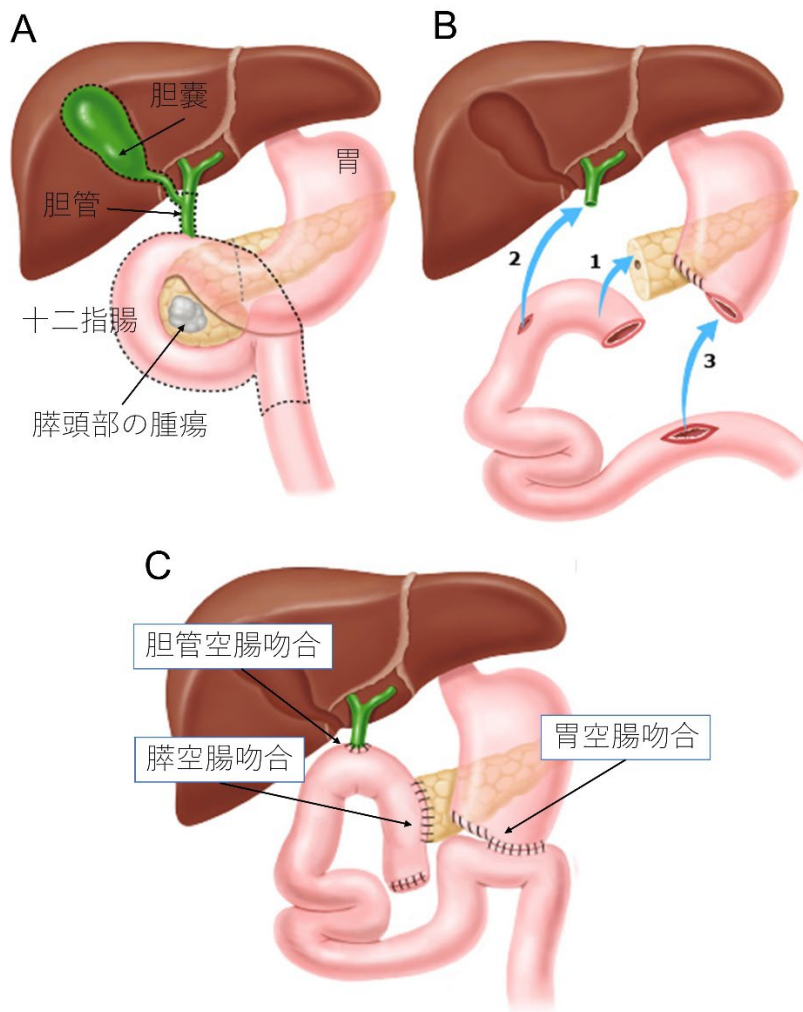


図1 PDの切除範囲と再建

A は膵頭部の腫瘍を切除する際の切除範囲である。胃の一部分と十二指腸、胆嚢を含む胆管、膵臓を一塊に切除する。B は腫瘍切除後の再建前の状態である。1 は膵臓の断端と空腸、2 は胆管の断端と空腸、3 は胃の断端と空腸をそれぞれ吻合し、C が再建の終了後の状態である (Howard et al., 2023 から引用し改変)。

<PDの合併症>

本術式の合併症発生率は35～51%と他の術式と比較し高いのが現状である。修練施設のPDの成績が良好とはいえ、約40%の合併症発生率との報告がある。合併症の内訳は膵液瘻、胃内容排泄遅延、術後出血があげられる (Harnoss et al., 2014; Okano et al., 2015; Pugalenti et al., 2016; Miura et al., 2016)。特に膵液瘻は13～20%の発生率であり、その治療のコントロールがつかぬ場合、仮性動脈瘤を形成し術後出血を招き致命的となる恐れがある (Callery et al., 2013)。仮性動脈瘤からの出血は総肝動脈からが多く、治療法として緊急開腹手術、動脈塞栓術があげられる (図2) (de Castro et al., 2005)。緊急対応できぬ場合、救命は困難である。急性期の救命が可能でも、のちに肝機能の低下を招き致命的となるケースもある。術後出血は膵液瘻を起因として発症するため、膵液瘻の低減が必要である。

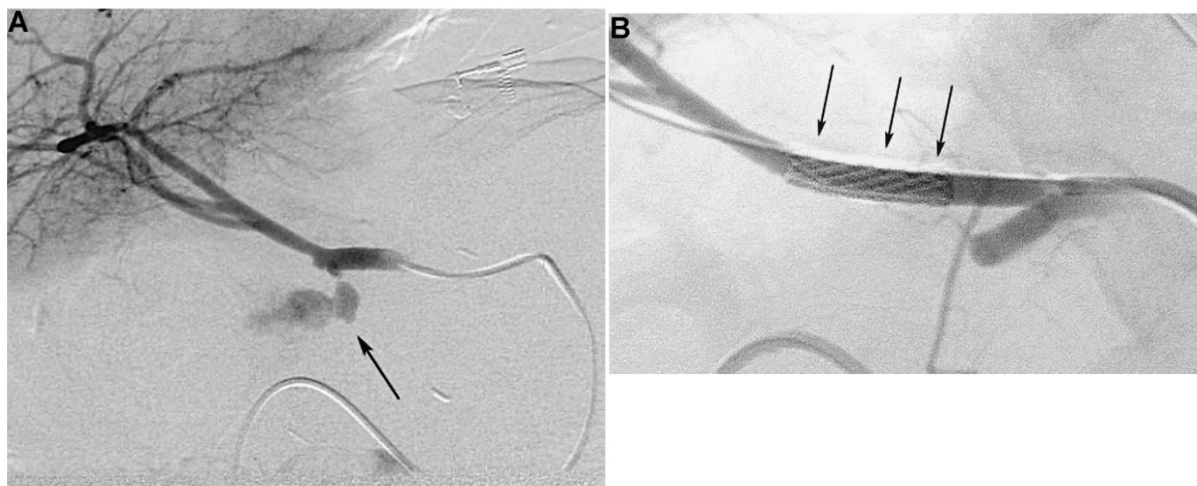


図2 仮性動脈瘤からの出血の治療

A 血管造影を行い、総肝動脈からの仮性瘤と出血を認める (矢印)。B 動脈塞栓術で矢印の範囲にステントが留置され、止血が得られている。(Castro et al., 2005 から引用し改変)。

<膵液瘻の評価>

膵液瘻は(International Study Group of Pancreas:以下 ISGP)に定義され、ドレーン排液のアミラーゼの数値から膵液瘻の有無を診断する。膵液瘻は Grade 分類され、臨床上の問題とならない Biochemical Leak と、治療介入が必要な Grade B、C に分類される (Bassi et al., 2017)。治療法はドレナージが最も効果的であり、抗生剤治療を要することもある。効果的な治療へつなげるためにも、患者や疾患の背景から膵液瘻リスクスコア (Fistula Risk Score: 以下 FRS) が作成された (Callery et al., 2013)。FRS の測定には膵臓の固さ、病理組織学的診断、主膵管の径、術中の出血量を用いて、それぞれ数値化して評価する (表 1)。出血量以外の項目は、術前から規定された因子であり、これらを術中に判断し術後管理を改善させることは膵液瘻の予防につながる。

表 1 Fistula Risk Score の測定項目

Risk factor	Parameter	Points
膵組織	硬	0
	軟	2
病理所見	膵癌もしくは膵炎	0
	乳頭部癌、十二指腸癌、膵嚢胞、膵内分泌腫瘍	1
主膵管の径 (mm)	≥ 5	0
	4	1
	3	2
	2	3
	≤ 1	4
術中出血量 (ml)	≤ 400	0
	401 -700	1
	701 -1000	2
	>1000	3

それぞれのポイントの合計で 1~2 points: Low risk, 3~6 points: Inter mediate, 7~10 points: High risk と評価される。(Callery et al., 2013 から引用し改変)。

<膵液瘻予防のための膵空腸吻合>

膵液瘻の発生は、膵空腸吻合部の縫合不全、膵液の漏出、また膵実質の損傷によって発症する。膵液瘻を予防するために膵空腸吻合の詳細な手技について考案されてきた (Grobyer et al., 2010) (図 3)。主な膵空腸吻合の手技は、2つのパートに分かれる。1つ目は膵管と消化管粘膜の吻合、2つ目は先の吻合部と膵断端を空腸で圧迫して覆う手技である。これらの手技は、1 cmに満たない数mm径の細い膵管と消化管の吻合を行い、膵損傷を避けるべく最小限の侵襲での膵実質へ操作を行う必要があり、高度な技術を要する。

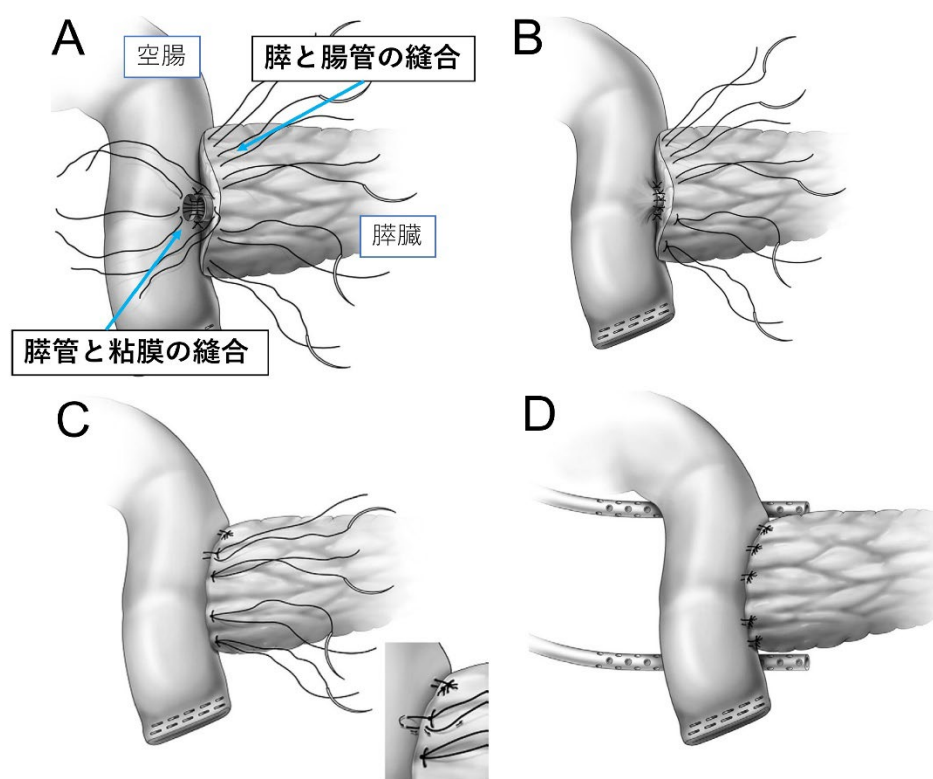


図 3 膵空腸吻合の手順

A 膵管と粘膜の縫合と膵と腸管の縫合をそれぞれ背側から行う。B 膵管と粘膜の結紮を行い膵管と粘膜の吻合が完了した状態。C 膵と腸管の縫合を行い、膵断端を腸管で覆う。D 再建の完了した状態。膵臓の上縁と下縁にドレーンを挿入し膵液瘻のモニタリングを行う。(Grobyer et al., 2010 から引用し改変)。

<外科医の技能評価と合併症>

膵液瘻の発症については、患者因子や膵空腸の手技だけでなく、外科医の手術技能との相関についても報告されている。

Birkmeyerらは、術中のビデオ評価スコアと患者の転帰を検討し、ビデオ評価スコアの低い術者群は有意に合併症の発症率が高いと報告した (Birkmeyer et al., 2013)。それまで外科医の手技と患者のアウトカムを検証した報告は極めて少なく、執刀医の術中パフォーマンスが術後成績を予測する可能性を示した点で、以降の外科教育分野の研究に大きな影響を与え、手術中の手技の評価について注目されるようになった。

近年のシステマティックレビューでは、手術技能評価スコアが高い外科医は術中術後合併症が少なく、術後成績が良好であることを示す研究が報告されている (Balvardi et al., 2022)。以上のことから外科医の技能は患者の予後へ影響する因子である可能性が示唆されてきた。

PDの領域においては手術中の膵空腸吻合を技能評価し、評価スコアが低い執刀群は術後の膵液瘻発症率が優位に高いという結果が報告された (Hogg et al., 2016)。Hggらの技能評価の方法は、膵空腸吻合を評価するための評価シート

Pancreaticojejunostomy score card (図4)を作成し、術中に撮影したビデオを評価した。

<技能向上のためのシミュレーショントレーニング>

外科医の技能向上のためには、様々な手術手技におけるシミュレーショントレーニングの有効性が報告されている (Stefanidis et al., 2015; Borgersen et al., 2018)。PD のトレーニングのためのシミュレーターも開発され、このシミュレーターでは膵空腸吻合、胆管空腸吻合、胃空腸吻合の再建の手技をトレーニングすることができ、汎用性が高い。複数回手技を施行することで手技に要する時間短縮や、評価スコアの向上など、その有用性が報告された (Tam et al., 2017)。手技を行う環境を 2D と 3D の映像で比較し、3D 映像の方が手技の時間が短縮し、評価スコアが向上したとの報告もある (Zwart et al., 2019)。



図5 PD のトレーニングのためのシミュレーター

PD の再建に必要な膵空腸吻合、胆管空腸吻合、胃空腸吻合の手技をそれぞれトレーニングに使用できる。(Tam et al., 2016 から引用し改変)。

PDは高度な技術を要する術式であり、技術習得には他の術式と比べ時間を要する。本邦において、PDの標準的な指導方法やトレーニングカリキュラムは存在せず、PDのシミュレーショントレーニング法も確立されていない。Tamら (Tam et al., 2017) はPDシミュレーターの使用環境、組織の再現性などトレーニングへ使用することの妥当性を証明したが、トレーニングカリキュラムの検証はされておらず、臨床上的成績との関係も評価していない。したがってPDシミュレーターは手術手技の評価に使用されているが、手術室でのパフォーマンスや臨床結果との相関についてはまだ報告されていない。

そこで、PDの技能評価スケールおよびシミュレーショントレーニングプログラムによる効率的な修練が、術後の成績を改善するのではないかと考えた。シミュレーターによる技能評価結果で手術室での成績や臨床転帰を予測することができれば、PDのような複雑な手術を外科医に行わせるための基準を設定することができ、今後の外科教育やトレーニングプログラムに活用することが可能となる。

本研究では、PDのトレーニングプログラムの作成へ向け、術後の合併症発生に大きく影響する膜空腸吻合をトレーニングするための技能評価スケールを開発し、教育的有効性を検討することを目的とした。また、手術室とシミュレーション環境下における膜空腸吻合の手技技能の相関関係を検討し、術前のシミュレーション環境下での評価結果と術者の手術室における技能や術後成績との相関を検討した。

方法

I. 膵空腸吻合評価スケール開発

外科教育の領域では、手術中の一般手技を評価する **Objective structured assessment of technical skill** (以下 **OSATS**) という評価スケールが開発され (Martin et al., 1997)、多くの研究及び手術教育の場で使用されてきた。一般手技だけでなく個別の術式に対する評価スケールも開発されており、PD に関してはロボット支援下手術における膵空腸吻合の技能評価スケールが Hogg らに開発された (Hogg et al., 2016)。この技能評価スケールは手術室で評価する際に使用するものであり、シミュレーション環境下での評価が困難である項目が含まれていたため、本研究においては手術室、シミュレーション環境下どちらにおいても評価可能なスケールを開発した。Hogg らのスケールを参考に膵空腸吻合の技能評価スケールの原案を作成し、共同研究者である肝胆膵外科高度技能専門医及び外科教育専門家の意見を加え統一したものを、膵空腸吻合技能評価スケール **Pancreaticojejunostomy performance assessment scale** (以下 **PJPAS**) (図 6) として本研究に使用した。評価項目内容と基準を詳細に検討するため評価基準表も作成した。

The pancreaticojejunostomy performance assessment scale

A. Pancreatic duct to mucoa sutures			Score
1	Needle handling	Suture bite and pitch to the duct	
2	Gentleness	No tear and slack in the duct	
3	Smooth	Suturing smoothly	
B. Knot tying pancreatic duct and jejunum			
4	Gentleness	No damage in the duct and jejunum	
5	Adaptation	Adaptation duct and jejunum	
6	Smooth	Knot tying smoothly	
C. Pancreatic pharencyma sutures			
7	Needle handling	Apporopreate stich position pancreas and jejunum	
8	Smooth	Suturing smoothly	
D. Knot tying pancreas and jejunum			
9	Adaptation	No gap pancreas and jejunum	
10	Smooth	Knot tying smoothly	
E. Suture management			
11	Suture management	Apporopreate order stures	
12	Assistant	Apporopreate use of assistant	
13	Smooth	Suture management smoothly	
F. Overall			
14	Final appearance	Covering pancreas by jejunum	
15	Smooth	Overall flow of procedure smoothly	

Good =2, Fair =1, Poor =0 total /30

図 6 Pancreaticojejunostomy performance assessment scale

本研究のために開発した膵空腸吻合の評価スケール

II. 膾空腸吻合の技能評価方法及び評価環境設定

a. 膾空腸吻合の評価方法

研究対象の外科医が執刀する PD の術野を手術室でビデオ撮影した。また、執刀の 1 か月以内に膾空腸吻合のシミュレーターを用いた手技をビデオ撮影した。北海道大学病院に所属する共同研究者の肝胆膵外科高度技能医 3 名が評価者となり、手術室環境のビデオ (Operating room video: 以下 ORV) とシミュレーション環境下のビデオ (Simulator video: 以下 SimV) を技能評価した。共同研究者の肝胆膵外科高度技能医はいずれも 100 例以上の PD 経験を有し、日本肝胆膵外科学会に認定された肝胆膵外科高度技能専門医であった。I で作成した膾空腸吻合技能評価スケール PJPAS と既存の OSATS を用いて技能評価を行った。対象となる外科医の背景情報、手術日、患者情報は評価者に提供しなかった。

b. 手術室環境での評価方法

手術室の天井の照明に搭載されたビデオカメラシステムを用いて術者の PD のビデオ撮影を行った (図 7)。ビデオカメラシステムが無い施設では、研究者が用意したビデオカメラを固定して撮影した。膾空腸吻合の手技が確実に撮影されるよう、研究者はすべての症例に立ち会った。評価対象となるビデオは、膾空腸吻合再建時の残膵もしくは挙上空腸に第 1 針目をかけるタイミングから評価開始として、再建が完成し最後の縫合糸を切るまでを評価終了とした。以上の手技の開始と終了が明確となるようビデオを編集し ORV とし、評価者が評価スケールに準じた項目を評価した。手術中の患者の安全を優先して、手術の難易度や術中の進行状況により、指導医が必要と判断した場合には術者を交代し、その症例を評価対象から除外した。本研究では明らかに指導医が途中から手技を行うビデオがあり、1 名を評価対象から除外した。

c. シミュレーション環境下における評価方法

1) シミュレーターの選定およびセッティング

国内で販売されている膵臓・空腸モデルから、より実際の手技に近い模擬吻合が可能な Biotexture™, (FASOTEC, Japan) を採用した (図 8)。このシミュレーターは汎用性に優れ、複数回使用しても評価や手技の難易度が変化しないモデルである。セッティングは、研究者が設計した木材を加工したフレームに、手術中に使用する実際のドレープ (3M™ ステリ・ドレープ™ 創傷保護用ドレープ) で覆い、その中にシミュレーターを設置し、手術室環境を再現した (図 9)。術者が手技を行う際に、同一条件で手技を行うことができるよう、シミュレーターの設置位置、手技を記録するビデオカメラの位置の標準位置を予め設定した。

2) シミュレーション環境下評価

上記の準備した枠内にシミュレーターを固定して、ビデオカメラを三脚で固定して術者のシミュレーターによる膣空腸吻合のビデオ撮影を行った（図 10）。研究者がすべての手技に立ち合い撮影条件を確認した。膣空腸吻合の再建の手順は各施設の手術室環境での手技に準じた方法で行った。持針器、撮子、縫合針や糸は研究者側で準備して運針、結紮の数など手術室環境と同じ手技となるように行った。膣もしくは空腸のシミュレーターに第 1 針目をかけるところから評価開始として、再建が完成し最後の糸を切るまでの手技を評価した。以上の手技の開始と終了が明確となるようビデオを編集し SimV とした。

3) シミュレーション環境についてのアンケート

ビデオ撮影の手技が終了後、対象の外科医にアンケートを実施した。シミュレーターとその手技を行う環境について、手術室環境と比較した評価の回答を依頼した。アンケートは膣臓を設置する位置の再現性、膣管径の再現性、各臓器の設置位置の再現性、手技に使用した器材の再現性、膣空腸吻合手技の難易度、トレーニングとして利用できるか、という項目に対して 5 段階評価で行った。また、実際の膣臓組織との硬度の比較を 5 段階で評価した。

4) パイロット試験による実施可能性の確認

上記設定条件のもと外科医の手技の評価の実施可能性についてパイロット試験を行った。PD 執刀 100 例以上の術者（PD エキスパート）と PD 執刀 10 例以下の外科医（PD 初心者）2 名の手技の撮影と評価を行い、評価スケール、シミュレーター、ビデオ記録のシステムが技能評価を行う際に問題がないことを確認した。ここで技能評価された 2 名の外科医は本研究対象者から除外した。



図7 手術室のビデオカメラシステム

ビデオカメラは術者側から清潔操作で動かすことが可能である。撮影場面の拡大縮小やフォーカスは術野外から調整可能である。

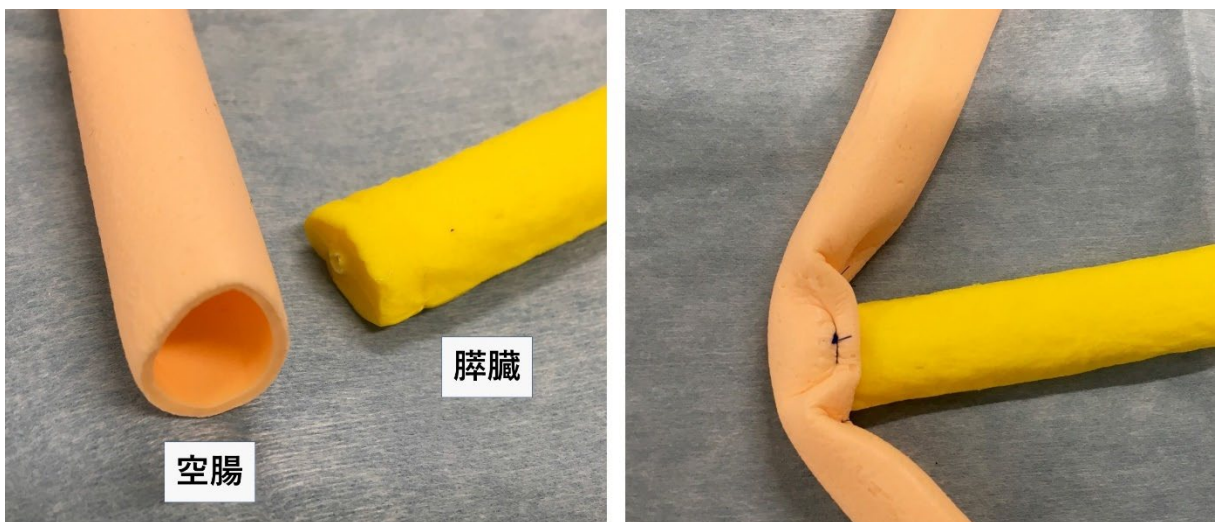


図8 膵臓と空腸のシミュレーター

膵臓モデルと空腸モデルは複数回使用することができる。右図は膵管空腸吻合の完了形である。

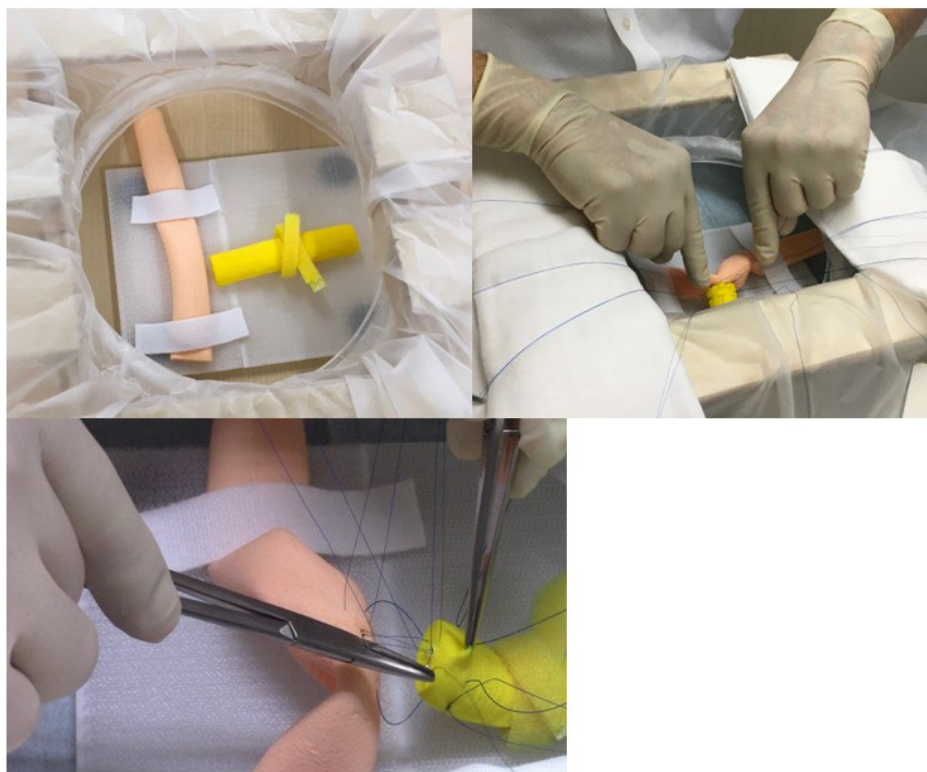


図9 シミュレーターの設定と手技

シミュレーターの固定位置を決め、常に同一条件で手技を行えるよう設定した（上段左図）。右
膝空腸吻合の結紮（上段右図）と膀胱に運針（下段）を行っている場面である。



図 10 シミュレーターのビデオ設置

図のような撮影条件となるよう位置を固定してビデオ撮影した。

III. 評価者トレーニング

評価者間の信頼性を担保するために外科教育専門家 2 名と評価者となる肝胆膵高度技能専門医 3 名で 2 時間の評価者トレーニングを行った。評価表の項目および評価基準を詳細に確認したのちに、テストビデオを用いて全ての局面、評価ポイントごとに評価方法の認識を統一した。

IV. 対象施設、対象者の選定

本研究は北海道内の肝胆膵外科修練施設もしくは、年間 PD 数が 10 例以上見込める施設を対象とした。研究参加者は 2018 年 3 月から 2020 年 3 月の間、対象施設で PD を執刀した外科医とした。外科医の経験年数は問わず、各施設で膵疾患、胆道疾患、十二指腸疾患のうち定型的な PD を施行した術式を対象とした。

V. 再建手技の違いについて

定型的な PD を行い再建方法が膵空腸吻合を行う症例を対象とした。施設や術者によって膵空腸吻合の運針の数や手順が異なる場合があるが、膵管と腸管の吻合、膵実質と腸管の吻合の手技が含まれていれば対象とした。シミュレーター環境下における手技は、実際の手術と同じ手技が行えるよう縫合糸を準備した。

VI. 評価項目と術後経過の記録

a. 主要評価項目と副次評価項目

本研究の主要評価項目は、手術室とシミュレーション環境下での膵空腸吻合技能評価スコアとの相関である。副次的評価項目は、評価スコアと術後成績の相関、および評価スコアと外科医の PD 経験との相関である。PD 経験については、執刀経験数によって Novice 群 (PD 執刀経験が 10 例以下)、Intermediate 群 (11 例以上、50 例以下)、Experienced 群 (51 例以上) の 3 群に分けて各項目を検討した。3 名の評価者の平均点を評価スコアとして用いた。

b. 術後経過の記録

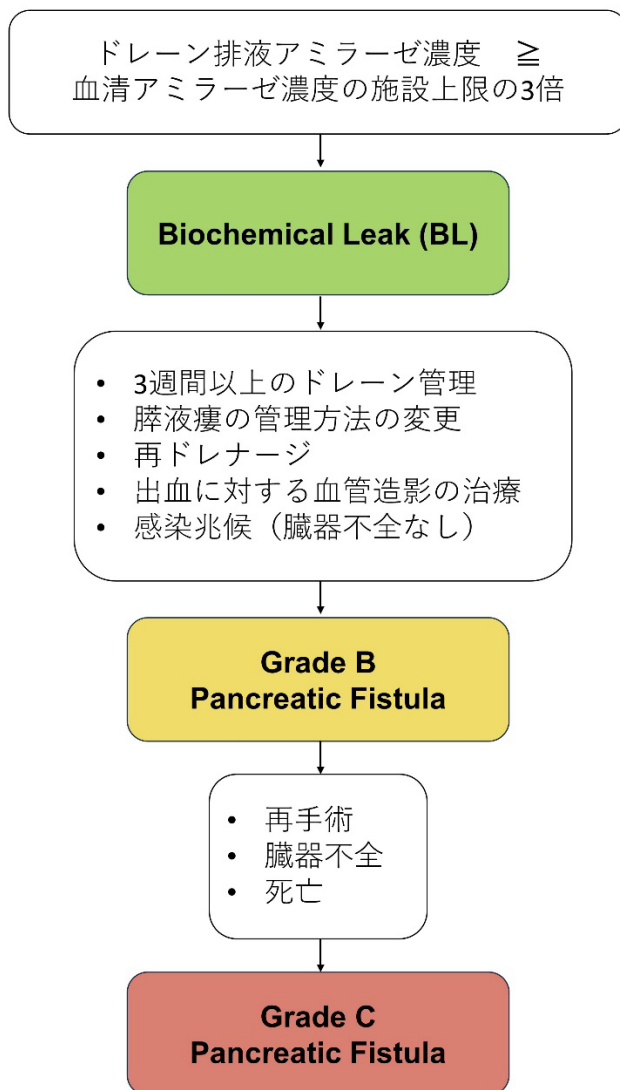
患者データは各参加施設のデータから、術後在院日数、術後合併症、その他の臨床データを記録した。患者の術後膵液瘻 (Postoperative pancreatic fistula、以下 POPF) の有無、および重篤な合併症の有無を記録した。術後合併症は Clavien-Dindo 分類を用いた (表 2)。その分類は I から V まで分類され、III と IV は a と b へ細分化される。分類の IIIb は全身麻酔の再手術を要する状態であり、IIIb 以上が重篤とされ V は死亡となる (Dindo et al., 2004)。膵液瘻は International Study Group の定義を用いて、ドレーン排液のアミラーゼの数値が血清アミラーゼの数値の 3 倍よりも高い場合に膵液瘻と診断した。また、図 11 のフローチャートで 3 つのグレードに分類

し、Grade C は臓器不全や再手術を要するような重篤な状態とされる(Bassi et al., 2017)。

表 2 Clavien-Dindo 分類

Grade	定義
Grade I	正常な術後経過からの逸脱があるが、薬物療法、外科的治療、内視鏡的治療、IVR 治療のいずれも必要としないもの。 ただし、制吐剤、解熱剤、鎮痛剤、利尿剤による治療、電解質補充、理学療法は上記治療には含めず、これらが行われていても Grade I とする。また、ベッドサイドでの創感染の開放は Grade I とする。
Grade II	制吐剤、解熱剤、鎮痛剤、利尿剤以外の薬物療法を要するもの。 輸血、中心静脈栄養を要する場合を含む。
Grade III	外科的治療、内視鏡的治療、IVR 治療を要するもの。
Grade IIIa	全身麻酔を要さない治療
Grade IIIb	全身麻酔下での治療
Grade IV	準集中治療室・集中治療室管理を要する生命を脅かす合併症。脳出血、脳梗塞、くも膜下出血などの中樞神経系の合併症を含む（ただし一過性脳虚血性発作は除く）。
Grade IVa	単一の臓器不全（透析を含む）
Grade IVb	多臓器不全
Grade V	患者の死亡
Suffix “d”	患者の退院時にも合併症が持続していた場合、接尾辞 “-d” (“disability”)を該当する合併症の Grade に付加する。

IVR: Interventional radiology (Dindo et al., 2004 から引用し和訳)。



いずれの項目も膵液瘻に起因するもの

図 11 膵液瘻の分類とフローチャート

Biochemical Leak は臨床症状なく、術後経過に影響を与えることは少ない。Grade B と Grade C の術後管理が重要となる。(Bassi et al., 2017 から引用し改変)。

VII. 統計学的手法

パイロット試験のデータから PD エキスパートと PD 初心者の手術室とシミュレーション環境下の合計得点率の差は 17%であった。 α 誤差 0.05、 β 誤差 0.20 でサンプルサイズ分析を行った結果、18名の参加者数が必要であることが示された。各施設の外科医のレベルが不明である点と、様々な要因による脱落の可能性を考慮し、本研究の目標サンプルサイズを 30名とした。

評価者間の相互信頼性は級内相関係数 (Intraclass correlation coefficients、以下 ICC) を用いて算出した。ORV と SimV の技能評価スコアの相関は Spearman の順位相関を用いて算出した。術後合併症の有無による二群間のスコアの比較には Mann-Whitney U 検定を用いた。膵液瘻の発症因子の検討に多変量解析を用いた。p 値が 0.05 未満を有意差ありとした。

統計解析には JMP 14 (SAS Institute Inc, Cary, NC, USA)及び SPSS version 26 (IBM, Armonk, NY, USA)を用いた。

VIII. 倫理的配慮

本研究は北海道大学の倫理委員会からの承認 (北海道大学臨床研究 IRB 承認申請番号 : 自 018-0194) に加え、研究参加施設倫理委員会の承認、および参加する外科医と患者からの書面同意を得て実施した。

結果

I. 膵空腸吻合評価スケール開発

a. PJ performance assessment scale

Hogg ら (Hogg et al., 2016)が報告した Pancreatojejunostomy scorecard を肝胆膵外科専門医の意見に基づいて修正し、本研究のために新たに Pancreaticojejunostomy performance assessment scale (以下 PJPAS) を作成した。この評価スケールは手術室とシミュレーション環境下の両方で技能評価可能な内容とした。評価スケールは膵管から粘膜への縫合と結紮、膵実質の縫合と結紮、縫合器材などの管理、総合評価の 6 つの主要項目に分け、さらに細部の項目を作成し全 15 項目の評価項目とした。得点内容は 0、1、2 点と評価尺度を使用し 30 点満点とした。この評価スケールは技能の評価だけでなく、トレーニングプログラムにおいて修練者に重点的にフィードバックすることを主として項目を選定した。Hogg らのスケール (図 4) では膵管へのドレナージチューブの固定の項目があったが、施設によってドレナージチューブの留置の有無や、チューブの太さの選択、留置の位置が異なるため評価項目から除外した。評価表を日本語に翻訳し、実際に使用した評価スケールは図 12、13 となる。評価スケールには評価項目の内容の記載があるが、評価者が評価採点の根拠を明確に定義する目的で評価基準表も合わせて作成した (図 14)。この評価基準表は評価者トレーニングの際にも全ての評価項目内容と評価基準の確認に使用した。

膵空腸吻合技能評価スケール

評価日 年 月 日 評価者名 ビデオ種類 Code

個別の手技評価 ○, △, ×の3段階評価

評価困難	①	膵管-腸管の運針	
<input type="checkbox"/>	<input type="checkbox"/>	1)	運針 針の湾曲に沿った運針ができていないか
<input type="checkbox"/>	<input type="checkbox"/>	2)	組織の取り扱い 組織に損傷をきたすような力が加わっていないか
<input type="checkbox"/>	<input type="checkbox"/>	3)	運針の円滑さ 運針の手技を円滑にできているか
	②	膵管-腸管の結紮	
<input type="checkbox"/>	<input type="checkbox"/>	4)	組織の損傷 結紮によって組織の損傷をおこしていないか
<input type="checkbox"/>	<input type="checkbox"/>	5)	膵管と粘膜の密着/出来栄 膵管と腸管が十分に密接しているか
<input type="checkbox"/>	<input type="checkbox"/>	6)	結紮の円滑さ 結紮の手技は正しく円滑にできているか
	③	膵実質と空腸の運針	
<input type="checkbox"/>	<input type="checkbox"/>	7)	運針 膵臓と腸管が密接する位置で運針できているか
<input type="checkbox"/>	<input type="checkbox"/>	8)	運針の円滑さ 運針の手技を正しく円滑にできているか
	④	膵実質と空腸の運針後の結紮	
<input type="checkbox"/>	<input type="checkbox"/>	9)	膵実質と空腸の密着 腸管と膵臓に隙間なく密接しているか
<input type="checkbox"/>	<input type="checkbox"/>	10)	結紮の円滑さ 結紮の手技を円滑にできているか
	⑤	糸の管理・助手との連携	
<input type="checkbox"/>	<input type="checkbox"/>	11)	糸の管理 複数の糸を順序よく管理できているか
<input type="checkbox"/>	<input type="checkbox"/>	12)	助手との連携・主体性 助手と連携できているか、主体的にできているか
<input type="checkbox"/>	<input type="checkbox"/>	13)	糸の管理の円滑さ 糸の管理を円滑にできているか
	⑥	最終形と全体の円滑さ	
<input type="checkbox"/>	<input type="checkbox"/>	14)	最終形 膵臓を腸管で覆った完成形を評価
<input type="checkbox"/>	<input type="checkbox"/>	15)	手技全体の円滑さ 総合的に円滑に手技はすすんでいるか

フリーコメント

図 12 技能評価シート (PJPAS)

PJPAS の内容を翻訳したもの。評価困難な項目があればチェックをいれる。

一般手技評価 OSATS: Objective assessment of technical skill

1～5点で評価

⑦ 組織への配慮	
1	頻回に組織に不必要な力が加わっている。不適切な道具の使用で組織の損傷がある
2	
3	組織に配慮した手技だが、時に不注意による組織の損傷がある
4	
5	一貫して適切な力で組織を扱い、組織損傷を最小限に留めている
⑧ 時間と動作	
1	多くの不必要な動きがある
2	
3	効率的な時間と動きであるが、時に不必要な動きがある
4	
5	明らかに、無駄のない動きで最大限の効率性がある
⑨ 器具の扱い方	
1	不適切な器具の使用によるためらいやぎこちない動作を繰り返している
2	
3	器具は適切に使用しているが 時にぎこちなさや固い動きがある
4	
5	流暢に器具を使用しぎこちなさがない
⑩ 手術の流れ	
1	よく手術が止まり、次の動作が不確かに見える
2	
3	ある程度適切な進行のための計画性が見られる
4	
5	明確な計画性により次の動作へ労力のない流れで進行している
⑪ 助手の利用 (今回は主体性も加味する)	
1	助手にアシストさせる場所が適切でない、またはうまく助手を使えていない (主体性なく終始指導を受けている)
2	
3	手術中はほぼ助手を使えている (時々助手の指導が入るが、主体性が見受けられる)
4	
5	手術中計画的に最大限助手を利用している (常に主体性を持っている)
⑫ 手技に特化した知識	
1	知識が欠如している。ほとんどの工程において具体的な指示を必要とする
2	
3	手術における重要な工程は理解している
4	
5	手術のすべての局面を熟知している

図 13 技能評価シート (OSATS)

OSATS の項目で評価する際に使用したもの

膝空腸吻合 評価基準

【評価全体についての共通認識】

膝空腸吻合は各施設の方法で行う。

3段階評価については大まかには以下の基準を参考に各項目の評価を行う。

高度技能医の審査ではなく、修練基準として評価するものである。

○	ほぼ理想的な手技 自立して手技が行えるレベル 臨床で患者に指導医なく実施して良いレベル エキスパートの理想的手技を100%とすると80%程度できていれば ○に該当
△	○でも×でもない 概ね手順は理解しているが手技に改善の余地があるレベル
×	よくない手技 手順の理解または手技全般に多くの改善が必要とされるレベル

○に該当しない場合、×であるかを判断し、×に該当しない場合に△とする。

【不完全なビデオに対する評価】

頭や手で視野が隠れ評価困難な場合であっても、観察できる範囲で評価を行う。

ただし、当該項目の手技が6割以上観察できない場合は、“評価困難項目”として評価表にチェックする（手元が6割以上隠れて評価困難 → 評価した上で+評価困難ボックスにチェック）。

①練習-空欄の選択		③講義質と空欄の選択		⑤糸の管理、助手との連携	
<p>選択の手技、選択時の視線への影響、選択の円滑さを評価する。</p> <ul style="list-style-type: none"> ● 全体的に選択のバイト、ピッチが理想的で、斜の角度に合った選択ができている。 ● 1針までは理想的でない手技があった場合でも○とする。 ● 選択の回転運動、斜し直しなど特定の選択方向で手技の改善が必要。 ● ○でもよいが理想的でない手技がある。 ● 全体の半分以上の手技に改善が必要 ● 斜の角度を無視した側入や側出、選択のバイト、ピッチがそろわないなど、粗共的な選択の仕方に関する指摘があり、練習が必要。 	<p>選択の手技と円滑にできているかを評価する。</p> <ul style="list-style-type: none"> ● 選択のバイト、ピッチ、斜の角度に合った選択ができている。 ● 時に選択のバイトやピッチがそろっていない。 ● 1針でも理想的でない選択操作がある。 	<p>1)選択</p> <ul style="list-style-type: none"> ● 斜の斜し直しが1針あったとしても、残りの選択が○に該当していれば△とする。斜の斜し直しが複数回ある場合は、×。 ● 視線を空想的に高い位置をきかしていない。 ● 膝球質をよける、優しく押すなどの動作に絞る必要がある。 ● 視線に視線をきかしていないが、速度変化がわかることがある。 ● Soft, pouchesの形例に適合する速度変化がわかる操作をしている。 ● 斜への圧縮、開化操作がわかる。 ● 膝管又は脚裏質に斜の斜し直し操作がある (1回)。 	<p>11)糸の管理</p> <ul style="list-style-type: none"> ● 複数の糸を順次よく管理できている。 ● 斜の糸の順序を把握して管理できている。 ● 1本もしくは1回の糸の順序の間違えを認めた。 ● 複数の糸を管理できていない。 ● 斜の順序を把握していない。 ● 複数回、糸の間違えを認めた。 <p>時間を超しても、確認しながら正しい糸の順序で糸を管理ができている場合は○とする。手解等も正しいが、斜の順序で糸が斜管に1回以上まわらなければ△とする。糸同士の間隔が狭まっておらず、全体の糸の順序に影響する場合は×とする。</p>	<p>○</p> <ul style="list-style-type: none"> ● 複数の糸を順次よく管理できている。 ● 斜の糸の順序を把握して管理できている。 ● 1本もしくは1回の糸の順序の間違えを認めた。 <p>△</p> <ul style="list-style-type: none"> ● 複数の糸を管理できていない。 ● 斜の順序を把握していない。 <p>×</p> <ul style="list-style-type: none"> ● 複数回、糸の間違えを認めた。 	<p>○</p> <ul style="list-style-type: none"> ● 複数の糸を順次よく管理できている。 ● 斜の糸の順序を把握して管理できている。 ● 1本もしくは1回の糸の順序の間違えを認めた。 <p>△</p> <ul style="list-style-type: none"> ● 複数の糸を管理できていない。 ● 斜の順序を把握していない。 <p>×</p> <ul style="list-style-type: none"> ● 複数回、糸の間違えを認めた。
<p>2)相織の</p> <ul style="list-style-type: none"> ● 速度変化による距離感等を認める。 ● 圧迫や減速時といった注意を要するような出血が複数回以上ある。 ● 液面が斜し直しがある。 	<p>7)選択</p> <ul style="list-style-type: none"> ● 選択のバイト、ピッチがそろわず、選択操作に粗共的な間隔がある。 <p>空間に全ての選択した後により直す場合は×とする。</p>	/		<p>○</p> <ul style="list-style-type: none"> ● 主目的の指が出しにより助手と連携できている。 <p>△</p> <ul style="list-style-type: none"> ● 時に主目的に欠ける。 ● 時に助手から指を受けられているが、副目的の指と連携できていない。 <p>×</p> <ul style="list-style-type: none"> ● 主目的の指が出しができておらず、助手と連携できていない。 ● 全体的に助手から指を受けられている。 <p>本研究ではシミュレーターでの助手は同一番で固定し、助手は指を受けられた時のみサポートする。</p>	<p>○</p> <ul style="list-style-type: none"> ● 主目的の指が出しにより助手と連携できている。 ● 時に主目的に欠ける。 ● 時に助手から指を受けられているが、副目的の指と連携できていない。 <p>△</p> <ul style="list-style-type: none"> ● 主目的の指が出しができておらず、助手と連携できていない。 ● 全体的に助手から指を受けられている。
<p>取り扱い</p> <ul style="list-style-type: none"> ● 距離をきたしたかどうかわからない×、前後の可能性があれば△。 ● 膝管状態例で膝子を膝管に入れて押したり、膝管チューブを用いて膝管を保持する操作が変更的であったり押したりよりよくなる場合は○となる。速度変化なく、斜の角度に合った選択が行われたにも関わらず出血が生じた場合、1針のみの出血であれば△、出血が複数回の選択で生じれば×と判断する。 	<p>8)選択の円滑さ</p> <ul style="list-style-type: none"> ● 工程を理解し、半順に沿って正しく円滑に手技が進んでいる。 ● 工程を理解し、半順に沿って手技が行えているが、時に円滑さにかける。 ● 工程を理解しておらず全体的に円滑に手技が進んでいない。 ● 半分以上のミスで円滑でない選択があった。 			<p>○</p> <ul style="list-style-type: none"> ● 工程を理解し、半順に沿って正しく円滑に手技が進んでいる。 <p>△</p> <ul style="list-style-type: none"> ● 工程を理解し、半順に沿って手技が行えているが、時に円滑さにかける。 <p>×</p> <ul style="list-style-type: none"> ● 工程を理解しておらず全体的に円滑に手技が進んでいない。 ● 半分以上のミスで円滑でない選択があった。 	<p>○</p> <ul style="list-style-type: none"> ● 工程を理解し、半順に沿って正しく円滑に糸の管理ができている。 ● 工程を理解し、半順に沿って糸の管理が行えているが、時に円滑さにかける。 ● 工程を理解しておらず、全体的に円滑に糸の管理ができている。 ● 半分以上のミスで糸の管理できていない。
<p>3)選択の円滑さ</p> <ul style="list-style-type: none"> ● 選択の工程を理解しておらず、全体的に円滑に手技が進んでいない。 ● 半分以上のミスで円滑でない選択があった。 <p>選択の工程とは斜を待つ、側入する、側出す動作、選択方向の順番などの手順までを含める。逆斜など斜の斜ら換えは正しい選択は斜管に2回以上まわらなければ○とする。①必要とされる斜の角度に対して、斜管に把持した角度が異なる斜管を要したり、②必要とされる斜の角度自体を理解していないように見えたり、③斜管に斜角度の調整を行うなど、誤行確認しているように判断される場合は×とする。</p>	<p>13)糸の管理の円滑さ</p>	<p>○</p> <ul style="list-style-type: none"> ● 工程を理解し、半順に沿って正しく円滑に糸の管理ができている。 ● 工程を理解し、半順に沿って糸の管理が行えているが、時に円滑さにかける。 ● 工程を理解しておらず、全体的に円滑に糸の管理ができている。 ● 半分以上のミスで糸の管理できていない。 			

②継管-空欄締結の結果 結果物の組織の粗り度、継管による密着、円滑にできているかを評価する。	④継管質と空欄の連結後の結果 結果物の強さ、円滑にできているかを評価する。	⑥最終形と全体の円滑さ
4) 組織の損傷 ○ 組織に空欄的粗りの粗り・継管により組織の損傷を起していない。 △ 組織を起し粗りがない様の粗り方をしている。 × 粗りもしくは空欄的粗りの損傷を起している。 ○ 継管もしくは空欄的粗りの損傷を起している。 △ 継管もしくは空欄的粗りの損傷を起している。 × 継管もしくは空欄的粗りの損傷を起している。		
5) 継管と構化管柱間の密着 ○ 継管が十分に密着している。 △ 継管が十分に密着しているが一部密着が不十分な結果を認める。 × 継管が不十分な結果を認める。	9) 継管の密着 ○ 継管が十分に密着している。 △ 継管が十分に密着しているが一部密着が不十分な結果を認める。 × 継管が不十分な結果を認める。	14) 最終形 ○ 継管は空欄で十分に覆われて理想的な最終形である。 △ 部分的に若干密着の余隙がある最終形である。 × 多くの密着を必要とする最終形である。
6) 結果の円滑さ ○ 工程を理解し、手順に沿って正しく円滑に手技が進んでいる。 △ 工程を理解し、手順に沿って手技が行えているが、時に円滑さや円滑さにかける。 × 工程を理解しておらず、全体的に円滑に手技が進んでいない。	10) 結果の円滑さ ○ 工程を理解し、手順に沿って正しく円滑に手技が進んでいる。 △ 工程を理解し、手順に沿って手技が行えているが、時に円滑さにかける。 × 工程を理解しておらず、全体的に円滑に手技が進んでいない。	15) 全体の円滑さ ○ 全工程において計画的に円滑に手技ができています。 △ 工程は計画的だが、全体を通して若干円滑さに欠ける。 × 全体的に計画的に欠け、円滑に手技が進んでいない。

図 14 PJPAS 評価基準表
3 段階評価の基準を明確にし、評価者が参考にできる評価基準表を作成した。

II. 膵空腸吻合評価の実行可能性の検討

a. パイロット試験結果

手術室及びシミュレーション環境下のビデオ撮影は問題なく実行可能であった。記録ビデオも評価するための画質に問題なく、評価者が項目毎に評価可能であった。

b. シミュレーターに対するアンケート結果

研究参加者からのアンケート結果を表3にまとめた。研究参加者からはこのシミュレーターが実際の手術を高く再現していると評価された。今後のトレーニングツールとしての可能性については全参加者から有効活用できるとの返答が得られた。

表3 研究参加者へのシミュレーターに対するアンケート結果 (N = 27)

シミュレーション環境の再現性	点数(1-5)
1. 膵臓の設置する位置	4 [3; 5]
2. 各臓器の設置する位置関係	4 [3.5; 5]
3. 使用する器材	5 [4; 5]
膵臓モデルの再現性	点数(1-5)
4. 膵臓の硬さ	3 [2; 3.5]
5. 主膵管の径 (mm)	4 [4; 5]
6. シミュレーターにおける手技の難易度	3 [3; 3]
7. 今後のトレーニングに有効に活用できるか	5 [5; 5]

1~3：再現性（1.再現できていない、3.ある程度再現できている、5.再現できている）。4：実際の膵臓と比較した硬度（1.soft pancreas、 3.medium、 5.Hard pancreas）。5：主膵管径（mm）。 6：難易度（1.とても簡単、3.普通、5.とても難しい）。7：今後のトレーニングに有効活用できるか（1.利用できない、3.わからない、5.利用できる）

III. 評価者間信頼性

ORV の 3 名の評価者評価結果から得られた級内相関係数を表 4 に示した。PJPAS を用いた評価では 3 名の評価者のどの組み合わせでも信頼性が低い結果となった。一方、OSATS を用いて評価した場合、評価者 B と C の組み合わせ以外では高い信頼性を認めた。本研究において ORV を PJPAS スコアで評価することは信頼性が担保されないと判断し、ORV の PJPAS スケール評価は後の研究結果解析から除外した。

表 4 手術室評価における評価者間信頼性検討 (N = 27)

	ICC (95% CI)	
	PJPAS score	OSATS score
Rater A vs rater B	0.17 (-0.24- 0.54)	0.55 (0.23-0.77)
Average	0.30 (-0.63-0.70)	0.71 (0.38-0.87)
Rater B vs rater C	0.02 (-0.38-0.42)	0.20 (-0.18-0.53)
Average	0.05 (-1.25-0.60)	0.34 (-0.43-0.69)
Rater A vs rater C	0.04 (-0.36-0.43)	0.51 (0.18-0.74)
Average	0.08 (-1.13-0.60)	0.68 (0.31-0.68)
Overall	0.15 (-0.09-0.44)	0.42 (0.19-0.65)
Average	0.35 (-0.31-0.70)	0.69 (0.42-0.85)

ICC: Intraclass correlation coefficients (95% CI: confidence interval), OSATS: Objective structured assessment of technical skill, PJPAS: Pancreaticojejunostomy performance assessment scale

SimV における評価者間の信頼性を表 5 にまとめた。PJPAS、OSATS どちらの評価スケールでも評価者間の信頼性は担保されていた。本研究において、シミュレーション環境下で評価する場合、PJPAS、OSATS どちらの評価スケールとも使用できることが示された。

表 5 シミュレーション環境下における評価者間信頼性検討 (N = 27)

	ICC (95% CI)	
	PJPAS score	OSATS score
Rater A vs rater B	0.44 (0.09-0.70)	0.44 (-0.23-0.49)
Average	0.61 (0.16-0.82)	0.62 (-0.61-0.66)
Rater B vs rater C	0.60 (0.28-0.79)	0.32 (-0.04-0.61)
Average	0.75 (0.44-0.88)	0.49 (-0.09-0.76)
Rater A vs rater C	0.56 (0.25-0.71)	0.52 (0.20-0.74)
Average	0.72 (0.40-0.87)	0.69 (0.33-0.85)
Overall	0.55 (0.32-0.74)	0.42 (0.19-0.65)
Average	0.78 (0.59-0.89)	0.69 (0.42-0.84)

ICC: Intraclass correlation coefficients (95% CI: confidence interval), OSATS: Objective structured assessment of technical skill, PJPAS: Pancreaticojejunostomy performance assessment scale

IV. 参加施設、参加者、患者背景

北海道内 11 施設から 30 名の外科医が本研究に参加した。そのうち、3 名の外科医がデータ解析対象から除外された。3 名の内訳は、手術ビデオの内容が不鮮明で評価が困難であった者が 1 名、手術手技を途中から明らかに指導医が行っていた者が 1 名、COVID-19 のパンデミック時でシミュレーターの手技が行えなかった者が 1 名であった。参加した施設及び外科医の背景を表 6、その外科医が執刀した患者背景を表 7、手術成績について表 8 にまとめた。

表 6 参加施設・参加者背景

参加施設数	11 施設
肝胆膵高度技能施設	4 施設
参加者数 (人)	27
卒後年次 (年)	17 [12; 23]
PD 執刀経験数	20 [4; 45]
直近 1 年の PD 執刀経験数	4 [2; 8]
全身麻酔下の執刀経験数	1 / 5 / 8 / 13
< 200 / 201-500 / 501-1000 / 1001<	

数値は中央値[四分位範囲]，PD: Pancreatoduodenectomy

表 7 患者背景

男性 / 女性	17 / 10
年齢	67 [57; 76]
BMI	22.9 [20.7; 25.5]
疾患名 (%)	
膵癌	14 (52%)
下部胆管癌	4 (15%)
十二指腸乳頭部癌	3 (11%)
膵管内乳頭粘液性腫瘍	2 (7%)
その他	4 (15%)

年齢及び BMI の数値は中央値 [四分位範囲]，BMI: Body Mass Index

表 8 手術成績

手術時間 (分)	514 [421- 632]
膵空腸吻合の手技時間 (分)	36 [30- 46]
術中出血量 (ml)	476 [305- 956]
Soft pancreas 症例数	15 (58%)
膵管径 (mm)	3 [2- 4]
術後 3 日目排液アミラーゼ値	399 [109- 1188]
入院期間 (日)	33 [18.3-36.5]
<hr/>	
膵液瘻数 (ISGPF 分類)	9 (33%)
なし	18 (66%)
Grade B	9 (33%)
Grade C	0 (0%)
<hr/>	
合併症 (Clavien-dindo 分類)	
なし	8 (30%)
Grade I	1 (4%)
Grade II	9 (33%)
Grade IIIa	9 (33%)
Grade IIIb	0 (0%)
<hr/>	
Fistula risk score	5 [4- 6]
No risk	2 (7%)
Low risk	3 (11%)
Intermediate risk	17 (63%)
High risk	5 (18%)

数値は中央値[四分位範囲]，症例数(%), ISGPF: International Study Group of Pancreas Fistula

V. 膵空腸吻合の評価と術後成績

a. 主要評価項目：手術室とシミュレーション環境下の技能評価スコア比較

手術室とシミュレーション環境下の OSATS によるビデオ技能評価スコアの相関を、Spearman の順位相関 r を用いて算出した結果、両環境下の技能評価スコアに相関性は認めなかった($r = 0.05, p = 0.81$) (図 15)。

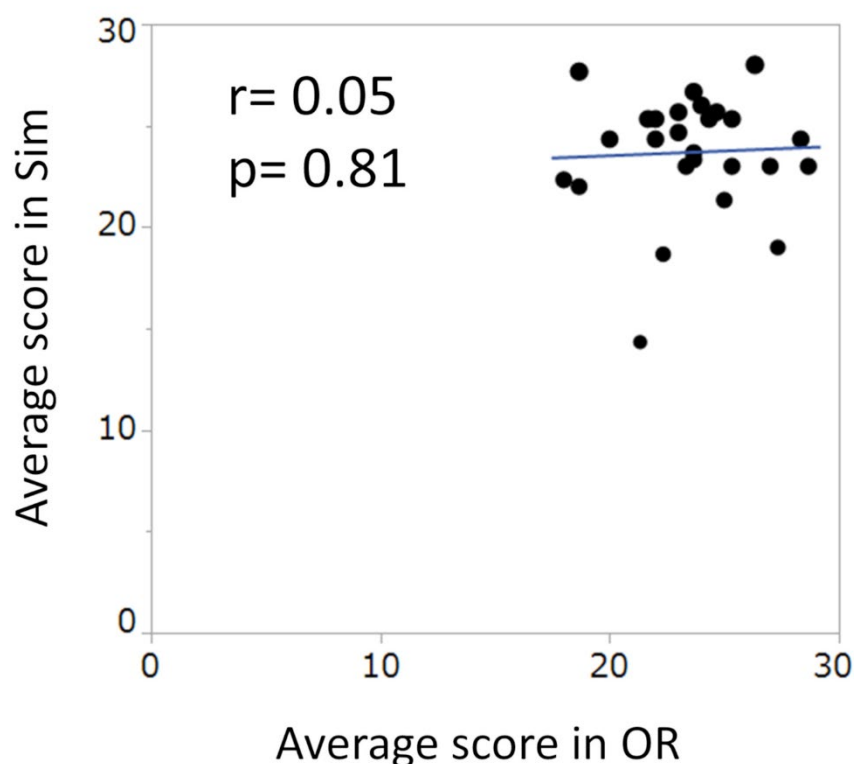


図 15 手術室とシミュレーターによる平均スコアの相関性

相関係数(r)は Spearman の順位相関を用いて算出した。OSATS で評価を行い、両群に相関性は認めなかった。

b. 副次評価項目：評価スコアと術後経過の検討

評価スコアと術後経過の解析をするにあたり、PD の執刀経験数によって Novice 群 (PD 執刀経験が 10 例以下)、Intermediate 群 (11 例以上、50 例以下)、Experienced 群 (51 例以上) の 3 群に分けた。それぞれの群における患者背景を表 9 に、3 群における評価スコア及び手術成績を表 10 に提示した。

表9 PD 経験数による分類と患者背景

	All (n = 27)	Novice (n = 10)	Intermediate (n = 11)	Experienced (n = 6)	P value
卒後年次	17 [12;23]	11 [7;12]	20 [16; 23]	26 [20; 32]	<0.01
直近1年PD執刀数	4 [2; 8]	1 [1; 3]	7 [4; 8]	15 [5; 30]	<0.01
高度技能専門医	8 (30%)	0 (0%)	3 (11%)	5 (19%)	<0.01
高度技能施設	19 (73%)	7 (26%)	7 (26%)	5 (19%)	0.73
患者背景					
性別 男	17 (63%)	7 (67%)	9 (75%)	2 (33%)	0.23
年齢 (歳)	67 [57;76]	69 [56;77]	72 [55;77]	63.5 [58;76]	0.89
Body mass index	22.9 [20.7; 25.5]	22.5 [21.1;25.2]	23 [19.6;24.4]	23.8 [20.4;29.2]	0.85
膵癌症例	14 (52%)	5 (56%)	6 (50%)	3 (50%)	0.96

数値は中央値[四分位範囲], 数(%), PD: Pancreatoduodenectomy, BMI: Body Mass Index

表10 PD 経験数による評価スコアと手術成績の相関検討

	All (n = 27)	Novice (n = 10)	Intermediate (n = 11)	Experienced (n = 6)	P value
Fistula Risk Score	5.0 [4.0; 6.0]	5.5 [3.5; 7.3]	4.0 [3; 6]	5.5 [3; 6]	0.63
OSATS score in OR	23.7[22.0; 25.3]	21.7 [18.7; 23.4]	25 [23; 27.3]	25.0[24.3; 26.8]	0.01
OSATS score in Sim	24.3[23.0; 25.3]	23.3 [21.2; 24.9]	23.3 [23.0; 25.3]	25.5[25.1; 26.5]	0.04
PJPAS score in Sim	24.0 [21.3; 26.0]	22.2 [20.3; 26]	23.7 [21.7; 25]	25.3[24.3; 28.5]	0.12
全手術時間 (分)	514 [421; 632]	574 [514; 677]	461 [380; 574]	442 [338; 603]	0.05
PJ 手技時間(分)	36 [30; 46]	46 [41; 49]	31 [28; 36]	30 [29; 37]	<0.01
Sim 手技時間 (分)	27.0 [25.0; 31.0]	31 [27; 36]	26 [23; 27]	29 [19; 31]	0.03
排液アミラーゼ 最高値	399 [109; 1188]	479 [202; 1352]	318 [109; 1178]	238 [24; 2770]	0.76
出血量 (ml)	476 [305; 956]	538 [380; 1275]	630 [202; 1910]	358 [269; 593]	0.41
術後合併症 CD III <	9 (33.3%)	3 (30%)	5 (45.4%)	1 (16.6%)	0.45
POPF (Grade B)	9 (33%)	4 (40%)	4 (36.3%)	1 (16.6%)	0.60

数値は中央値[四分位範囲], 数(%), CD III: Clavien-dindo 分類III, OSATS: Objective structured assessment of technical skill, PJPAS: Pancreaticojejunostomy performance assessment scale, OR: Operating room (手術室), Sim: Simulation (シミュレーション), POPF: Postoperative pancreatic fistula, PJ: Pancreaticojejunostomy (膵空腸吻合)

POPFの有無、および重篤な合併症の有無について2群に分け、それぞれの群のスコアについて解析した(表11)。POPFを発症した群は、発症しなかった群に比べて、SimVのPJASスコアが低い傾向があった。同様にClavian-Dindo分類のGradeIII以上の術後合併症を発症した群は、合併症のない群に比べORVでのPJスコアは低い傾向にあった。

表11 術後合併症発症の有無と評価スコアの比較

	膵液瘻なし (n = 18)	膵液瘻あり (n = 9)	P value
OSATS score in OR	23.5 [21.6; 25.1]	24.3 [22.2; 27.7]	0.95
PJPAS score in Sim	25.2 [21.8; 27.1]	23.0 [20.5; 24.2]	0.05
OSATS score in Sim	25.0 [22.8; 25.8]	23.0 [22.7; 24.3]	0.10
	CD Grade <III (n = 17)	CD Grade >III (n = 10)	
OSATS score in OR	23.3 [21.5; 24.8]	24.5 [22.0; 28.3]	0.20
PJPAS score in Sim	25.0 [21.0; 26.7]	23.5 [21.6; 24.4]	0.33
OSATS score in Sim	24.7 [21.7; 25.7]	24.0 [23.0; 24.6]	0.66

数値は中央値[四分位範囲], 数(%), OSATS: Objective structured assessment of technical skill, PJPAS: Pancreaticojejunostomy performance assessment scale, OR: Operating room (手術室), Sim: Simulation (シミュレーション)

膵液瘻の発症について術後結果と評価スコアを検討因子として単変量、多変量解析を表12に提示する。多変量解析によると、膵液瘻の発症因子は、既存のFistula risk scoreに加えてシミュレーターにおけるPJPASスコアの低値が影響しうる因子と考えられた。また、シミュレーション環境下によるOSATSおよびPJPAS評価スコアと術者因子の相関に関する解析結果を表13に提示する。シミュレーション環境下においてOSATS及びPJPAS評価でPDの執刀経験数と中等度の相関を認めた。

表 12 膵液瘻の発症因子の検討

	膵液瘻なし (n = 18)	膵液瘻あり (n = 9)	P value	多変量解析
OSATS score in OR	23.5 [21.6; 25.1]	24.3 [22.2; 27.7]	0.95	
PJPAS score in Sim	25.2 [21.8; 27.1]	23.0 [20.5; 24.2]	0.05	0.09
OSATS score in Sim	25.0 [22.8; 25.8]	23.0 [22.7; 24.3]	0.10	0.44
卒後年次	17 [11.5; 23]	16 [13.5; 24]	0.59	
PD 経験	30 [3.8; 60]	20 [3.5; 40]	0.66	
肝胆膵高度技能医	5 (28%)	3 (33%)	0.77	
全手術時間	511 [412; 667]	515 [430; 588]	0.84	
PJ の手技時間	36 [30; 47]	35 [26; 44]	0.35	
Sim の手技時間	28 [26; 30]	26 [22; 34]	0.78	
患者年齢	69 [58; 76]	66 [57; 77]	0.90	
BMI	22.1 [19.6; 26]	23.2 [22.5; 24.9]	0.38	
Fistula risk score	4 [1.8; 6]	6 [4.5; 7.5]	0.05	0.04

数値は中央値[四分位範囲], 数(%), PD: Pancreatoduodenectomy, BMI: Body Mass Index, OSATS: Objective structured assessment of technical skill, PJPAS: Pancreaticojejunostomy performance assessment scale, OR: Operating room (手術室), Sim: Simulation (シミュレーション), PJ: Pancreaticojejunostomy (膵空腸吻合)

表 13 シミュレーション環境下による評価と術者因子の相関関係

OSATS score との比較	r	P value
卒後年次	0.45	0.02
PD 経験	0.52	0.005
Sim の手技時間	-0.14	0.47
PJPAS score との比較		
卒後年次	0.30	0.14
PD 経験	0.45	0.02
Sim の手技時間	-0.07	0.73

OSATS: Objective structured assessment of technical skill, PJPAS: Pancreaticojejunostomy performance assessment scale, PD: Pancreatoduodenectomy, Sim: Simulation (シミュレーション)

考察

本研究で手術室とシミュレーション環境下における膣空腸吻合技能の評価を行った結果、2つの環境下における評価結果の相関性を認めなかった。膣空腸吻合の手術成績の評価と患者の臨床転帰の分析では、シミュレーション環境下の PJPAS スコアは患者の POPF 発生率と相関を示した。シミュレーション環境下における OSATS、PJPAS それぞれの評価は PD の執刀経験数と中等度の相関を認め、PD 習熟度の指標となる可能性がある。これらはシミュレーション環境を利用して術者の技能を手術前にある程度予測することが可能で、手術前の到達目標の基準やトレーニングとして、有効活用できると考えられる。膣臓の状態や患者の因子によって実際の手術の難易度は大きく異なるため、客観性をもって手術室における執刀医のパフォーマンスを評価することは今後の課題である。

<評価スケール PJPAS 開発>

膣空腸吻合の評価表の作成は既存の評価表を参考に、研究者と高度技能専門医と議論を重ね、手術室環境とシミュレーター環境下において評価可能となるものを作成した。近年、医学領域のガイドライン作成や治療指針コンセンサス決定、さらに汎用性を求める技能評価スケール開発手法に、専門家への無記名アンケートと意見集約を繰り返し行い、コンセンサスを形成するデルファイ法を採用する例が増えている。実際に既存の外科系技能評価スケール開発にもデルファイ法の採用が報告されている (Zevin et al., 2013; Kurashima et al., 2019)。本研究では既存の研究で開発され、臨床成績との相関が証明された Hogg らの (Hogg et al., 2016) 評価スケールを本研究用に修正したこと、また、以前より広く外科教育領域で使用されている OSATS を採用したため、デルファイ法のプロセスを経ることなく評価スケールを作成した。

<評価者間信頼性と評価者トレーニング>

本研究で予想された課題の一つは、高難度手技中でも特に繊細な手技が求められる膣空腸吻合手技評価において、いかに評価者間の信頼性を担保するかという点であった。膣空腸吻合の手技の時間の中央値は手術室では 36 分、シミュレーション環境下では 27 分であり、30 名の参加者の両環境下におけるビデオ技能評価には約 30 時間を要した。評価者トレーニングを事前準備も含めて約 2 時間行ったにも関わらず、十分な評価者間信頼性が得られなかった。手術室における膣空腸吻合の PJPAS 評価結果は、高度な手技を詳細な特異的評価スケールで評価することの困難さを示唆した。級内相関係数は、評価ツールの信頼性を検証するために単一尺度と平均尺度の 2 つの尺度を用いる。単一尺度は、複数の項目のうちどれか一つを使用し、ツールの信

信頼性を計算する。平均尺度は複数の評価の平均値を使用し信頼性を計算する方法で、一般に単一尺度より優れているとされる (Koo and Li., 2016)。本研究のデータでは3人の評価者の平均点を用いて計算し、級内相関係数は ORV では PJPAS 評価で 0.35、OSATS 評価で 0.78 であった。また SimV では PJPAS 評価が 0.69、OSATS 評価が 0.69 であった。既存の Poudel ら (Poudel et al., 2016)、Kurashima ら (Kurashima et al., 2022)の研究で開発された腹腔鏡下鼠径ヘルニア修復術および腹腔鏡下胃切除術技能評価スケールの信頼性検証では、いずれも級内相関係数は 0.7 より高値であったと報告されている。しかし、Hogg らによる膵頭十二指腸切除術技能評価の研究では、2名の評価者間の信頼性は 0.15 から 0.4 であった (Hogg et al., 2016)。また、Tam ら (Tam et al., 2017)の研究では膵頭十二指腸切除術のシミュレーターを使用し OSATS を使用した評価者間の信頼性は級内相関係数が 0.67 という中等度の結果であった。既にエビデンスが確立された評価ツールである OSATS を使用しても、より高度な技術を要する肝胆膵外科領域の評価については評価者間信頼性の確保が難しいことが示唆された。Gawad らは評価者間の信頼性を高めるための適切な評価者数や時間を検討したが、十分なエビデンスがないのが現状である (Gawad et al., 2019)。

手術手技の評価において、評価者間の信頼性を向上させるための評価者トレーニングが推奨されている (Robertson et al., 2018)。本研究では評価者トレーニングを十分行ったにも関わらず、シミュレーターの評価環境で良好な信頼性が得られたものの、手術室の環境下での評価において高い信頼性は得られなかった。本研究の評価者は、肝胆膵外科のエキスパートであり、外科医のパフォーマンス評価に習熟した外科教育の専門家ではなかった。しかし、臨床の場における肝胆膵外科領域の手術技能評価や指導を考慮した場合、外科教育を専門としない肝胆膵外科の指導医が術中指導や技能評価、フィードバックを行うのが通常である。したがって本研究の評価スコアは日常の手術教育の環境を反映した結果に近いものであると考えられる。

本研究では、シミュレーション環境下での評価者間信頼性は PJPAS スコアと OSATS スコアともに高い結果であった。Tam ら (Tam et al., 2017)も同様にシミュレーション環境下における良い評価者間信頼性を報告している。シミュレーション環境は、患者因子、時間制限や術中リスク、手術中の助手やスタッフを含めた手術室環境因子など、実際の手術に影響を与えうる様々な交絡因子がないため、一律の条件下で手技を行うことが可能である。この点が評価者間の信頼性確保に寄与していたと考えられる。

<技能評価スコアと術後成績>

本研究では、シミュレーション環境下による技能評価スコアと、同一術者が執刀した手術の術後膵液瘻の発症に相関があることが示された。

本研究においてシミュレーション環境下で PJPAS 評価スコア 25 点以上の外科医の執刀症例では術後臍液瘻を発症せず、術後合併症も少ない結果であった。シミュレーション環境下における PJPAS 評価スコアは POPF 発生と相関を示した一方で、手術室とシミュレーション環境下における技能評価スコアの相関性は見出せなかった。その理由として、手術室における技能評価の評価者間信頼性が低いことが要因の一つと考えられた。手術室では患者や術者の様々な交絡因子が含まれるため、正確な手術室における評価ができていない可能性がある。その点で、シミュレーション環境下の評価は、術者の実際の技能レベルを反映している可能性がある。

2つの環境の評価の相関関係を検討した図 15 によると、得点分布に中間層や明らかな低得点者が存在しなかった。他の得点者が存在しただけで相関性を認める可能性もある。本研究は PD の執刀の後にシミュレーターの手技を行う設定であったため、PD 執刀を許可された者のみが対象となった。PD 経験が少ない外科医であっても外科医としての経験が担保されており、低得点者が存在しなかった可能性がある。PD 未経験で、これからトレーニングを要する者のレベルは様々であることが予想される。

PD のラーニングカーブは 40-80 症例と報告されている (Tseng et al., 2007; Fisher et al., 2012; Boone et al., 2015)。加えて、PD シミュレーターの Biotissue モデルを用いた研究では臍空腸吻合、胆管空腸吻合、胃空腸吻合の 3 つの手技を比較し、臍空腸吻合が最も長いラーニングカーブを要したと報告されている (Tam et al., 2017)。したがって、臍空腸吻合は非常に難度が高い手術手技であり、外科医がその手技を習得するためには時間と経験に加え、効率的なトレーニングが求められる。

本研究ではさらにシミュレーション環境下の技能評価スコアは PD の執刀経験数と相関を認めた。このことは、シミュレーション環境下の技能評価は PD 執刀経験レベルを予測する尺度として使用できる可能性を示唆している。本研究では PD を執刀した経験のある外科医のみが研究対象者であったため、手術室における PD 初回執刀の技能評価データはない。将来的にはシミュレーション環境下の技能評価が手術室での手術執刀の条件になりうるか、またエキスパートと同様のスコアを有するまでシミュレーショントレーニングをした外科医の初回術後成績の検討や、ラーニングカーブの短縮を検討する研究が必要と考えられる。

<研究の限界>

本研究には以下の限界がある。一つ目は手術室で手術技能評価を行う症例の背景がコントロールされていないことである。実際の手術中の手技の難易度は、患者因子、臍臓実質の状態や手術室環境などによって変化しうる。このような患者要因のバリエーションが手術室における外科医の技能評価結果に影響した可能性がある。二つ目は、

手術室ビデオの撮影環境が施設によって異なる点である。事前に撮影条件を確認し、術中ビデオ撮影に立ち会うことでこの差異を最小限にするよう努め、ビデオ評価時に評価困難な症例は除外した。そのような条件であっても実際にビデオを評価する際には評価対象となるタイミングにおいて、膵臓および空腸を観察する角度や拡大の程度に若干の違いが評価結果に影響した可能性がある。

総括および結論

1) 本研究から得られた新知見

1. 膾空腸吻合技能を手術室、シミュレーション両方の環境下で評価可能なスケールを開発した。
2. 膾空腸吻合の評価を行うシミュレーター環境を設定し、研究参加者よりトレーニングへ有効活用を期待する評価が得られた。
3. 手術室環境とシミュレーション環境下の同一術者の手術技能評価スコアに相関を認めなかった。高難度手技である膾空腸吻合の手術室における評価法にはさらなる検討の必要がある。
4. シミュレーション環境下の技能評価スコアは、同一術者が執刀した患者のアウトカムと相関を認めた。

2) 新知見の意義

特定の外科手術の手技技能を実際の手術で初めて評価するのではなく、事前にシミュレーション環境下で測定することができれば、術者の対象術式への習熟度の予測が可能となり、また、不足した習熟度をシミュレーショントレーニングで補填できる可能性がある。そのシミュレーショントレーニングにおいて評価スコアは達成目標のマイルストーンとなり得る。さらに、評価方法や環境の整備、データの蓄積で、手術技能評価により患者のアウトカムを予測できる可能性がある。

3) 今後の展開

膾頭十二指腸切除術は修練医にとって執刀機会が少なく、特に難度の高い膾空腸吻合などの手技には手術室外におけるトレーニングも必要である。今後の研究では、膾空腸吻合のシミュレーショントレーニングのカリキュラムを開発実施し、シミュレーショントレーニングの効果を評価することで、手術執刀指標の確立や臨床成績の向上効果に関する研究を検討したい。

謝辞

本稿を終えるにあたり、本研究の機会を与えていただき、終始多大なるご指導を頂いた北海道大学大学院医学研究院消化器外科学教室Ⅱ教授 平野 聡先生に深謝申し上げます。研究の企画、実施、科学的考察と本稿の執筆にあたり多くの助言、指導をいただいた北海道大学大学院医学研究院クリニカルシミュレーションセンター准教授 倉島 庸先生に深い感謝の意を表す。本研究の基礎論文の執筆に多くの指導をいただいた北海道大学大学院医学研究院消化器外科教室Ⅱ サシーム パウデル先生、研究を進めるにあたり多くの助言と指導をいただいた北海道大学病院 医療・ヘルスサイエンス研究開発機構 渡邊 祐介先生、手術ビデオ評価に多大なる時間をかけていただいた 北海道大学大学院医学研究院消化器外科学教室Ⅱ 岡村 圭祐先生、中村 透先生、野路 武寛先生に深い感謝の意を表す。最後に本研究へご協力いただいた全ての参加者の皆様に心より感謝申し上げます。

利益相反

本研究において開示すべき利益相反状態はない。

引用文献

Balvardi, S., Kammili, A., Hanson, M., Mueller, C., Vassiliou, M., Lee, L., Schwartzman, K., Fiore, J.F., Jr., and Feldman, L.S. (2022). The association between video-based assessment of intraoperative technical performance and patient outcomes: a systematic review. *Surg. Endosc.* 36, 7938-7948.

Bassi, C., Marchegiani, G., Dervenis, C., Sarr, M., Abu Hilal, M., Adham, M., Allen, P., Andersson, R., Asbun, H.J., Besselink, M.G., et al. (2017). The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 Years After. *Surgery.* 161, 584-591.

Birkmeyer, J.D., Finks, J.F., O'Reilly, A., Oerline, M., Carlin, A.M., Nunn, A.R., Dimick, J., Banerjee, M., Birkmeyer, N.J., and Michigan Bariatric Surgery, C. (2013). Surgical skill and complication rates after bariatric surgery. *N. Engl. J. Med.* 369, 1434-1442.

Boone, B.A., Zenati, M., Hogg, M.E., Steve, J., Moser, A.J., Bartlett, D.L., Zeh, H.J., and Zureikat, A.H. (2015). Assessment of quality outcomes for robotic pancreaticoduodenectomy: identification of the learning curve. *JAMA Surg.* 150, 416-422.

Borgersen, N.J., Naur, T.M.H., Sorensen, S.M.D., Bjerrum, F., Konge, L., Subhi, Y., and Thomsen, A.S.S. (2018). Gathering Validity Evidence for Surgical Simulation: A Systematic Review. *Ann Surg.* 267, 1063-1068.

Callery, M.P., Pratt, W.B., Kent, T.S., Chaikof, E.L., and Vollmer, C.M., Jr. (2013). A prospectively validated clinical risk score accurately predicts pancreatic fistula after pancreatoduodenectomy. *J Am Coll Surg.* 216, 1-14.

de Castro, S.M., Kuhlmann, K.F., Busch, O.R., van Delden, O.M., Lameris, J.S., van Gulik, T.M., Obertop, H., and Gouma, D.J. (2005). Delayed massive hemorrhage after pancreatic and biliary surgery: embolization or surgery? *Ann Surg.* 241, 85-91.

Dindo, D., Demartines, N., and Clavien, P.A. (2004). Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg.* 240, 205-213.

Fisher, W. E., Hodges, S. E., Wu, M. F., Hilsenbeck, S. G., and Brunicaudi, F. C. (2012). Assessment of the learning curve for pancreaticoduodenectomy. *Am J Surg.* 203, 684-690

Gawad, N., Fowler, A., Mimeault, R., and Raiche, I. (2019). The Inter-Rater Reliability of Technical Skills Assessment and Retention of Rater Training. *J Surg Educ.* 76, 1088-1093.

Grobmyer, S.R., Kooby, D., Blumgart, L.H., and Hochwald, S.N. (2010). Novel pancreaticojejunostomy with a low rate of anastomotic failure-related complications. *J Am Coll Surg.* 210, 54-59.

Harnoss, J.C., Ulrich, A.B., Harnoss, J.M., Diener, M.K., Buchler, M.W., and Welsch, T. (2014). Use and results of consensus definitions in pancreatic surgery: a systematic review. *Surgery.* 155, 47-57.

Hogg, M.E., Zenati, M., Novak, S., Chen, Y., Jun, Y., Steve, J., Kowalsky, S.J., Bartlett, D.L., Zureikat, A.H., and Zeh, H.J., 3rd (2016). Grading of Surgeon Technical Performance Predicts Postoperative Pancreatic Fistula for Pancreaticoduodenectomy Independent of Patient-related Variables. *Ann Surg.* 264, 482-491.

Howard A Reber. Surgical resection of lesions of the head of the pancreas. Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. <http://www.uptodate.com>. (Accessed on May 26, 2023.)

Koo, T.K., and Li, M.Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine.* 15, 155-163.

Kurashima, Y., Kitagami, H., Teramura, K., Poudel, S., Ebihara, Y., Inaki, N., Nakamura, F., Misawa, K., Shibao, K., Nagai, E., et al. (2022). Validation study of a skill assessment tool for education and outcome prediction of laparoscopic distal gastrectomy. *Surg Endosc.* 36, 8807-8816.

Martin, J.A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C., and Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 84, 273-278.

Miura, F., Yamamoto, M., Gotoh, M., Konno, H., Fujimoto, J., Yanaga, K., Kokudo, N., Yamaue, H., Wakabayashi, G., Seto, Y., et al. (2016). Validation of the board certification system for expert surgeons (hepato-biliary-pancreatic field) using the data of the National Clinical Database of Japan: part 2 - Pancreatoduodenectomy. *J Hepatobiliary Pancreat Sci.* 23, 353-363.

Okano, K., Hirao, T., Unno, M., Fujii, T., Yoshitomi, H., Suzuki, S., Satoi, S., Takahashi, S., Kainuma, O., and Suzuki, Y. (2015). Postoperative infectious complications after pancreatic resection. *Br J Surg.* 102, 1551-1560.

Poudel, S., Kurashima, Y., Kawarada, Y., Watanabe, Y., Murakami, Y., Matsumura, Y., Kato, H., Miyazaki, K., Shichinohe, T., and Hirano, S. (2016). Development and validation of a checklist for assessing recorded performance of laparoscopic inguinal hernia repair. *Am J Surg.* 212, 468-474.

Pugalethi, A., Protic, M., Gonen, M., Kingham, T.P., Angelica, M.I., Dematteo, R.P., Fong, Y., Jarnagin, W.R., and Allen, P.J. (2016). Postoperative complications and overall survival after pancreaticoduodenectomy for pancreatic ductal adenocarcinoma. *J Surg Oncol.* 113, 188-193.

Robertson, R.L., Vergis, A., Gillman, L.M., and Park, J. (2018). Effect of rater training on the reliability of technical skill assessments: a randomized controlled trial. *Can J Surg.* 61, 15917.

Stefanidis, D., Sevdalis, N., Paige, J., Zevin, B., Aggarwal, R., Grantcharov, T., Jones, D.B., and Association for Surgical Education Simulation, C. (2015). Simulation in surgery: what's needed next? *Ann Surg.* 261, 846-853.

Tam, V., Zenati, M., Novak, S., Chen, Y., Zureikat, A.H., Zeh, H.J., 3rd, and Hogg, M.E. (2017). Robotic Pancreatoduodenectomy Biotissue Curriculum has Validity and Improves Technical Performance for Surgical Oncology Fellows. *J Surg Educ.* 74, 1057-1065.

Tseng, J.F., Pisters, P.W., Lee, J.E., Wang, H., Gomez, H.F., Sun, C.C., and Evans, D.B. (2007). The learning curve in pancreatic surgery. *Surgery.* 141, 694-701.

Whipple, A.O., Parsons, W.B., and Mullins, C.R. (1935). TREATMENT OF CARCINOMA OF

THE AMPULLA OF VATER. *Ann Surg.* 102, 763-779.

Zevin, B., Bonrath, E.M., Aggarwal, R., Dedy, N.J., Ahmed, N., Grantcharov, T.P., and group, A. (2013). Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg.* 216, 955-965.

Zwart, M.J.W., Fuente, I., Hilst, J., de Rooij, T., van Dieren, S., van Rijssen, L.B., Schijven, M.P., Busch, O.R.C., Luyer, M.D., Lips, D.J., et al. (2019). Added value of 3D-vision during laparoscopic biliary pancreatico- and hepaticojejunostomy (LAELAPS 3D2D): an international randomized cross-over trial. *HPB.* 21, 1087-1094.