



Title	造影超音波検査による早期乳癌の腋窩リンパ節転移診断に関する研究：人工知能併用および非併用での検証
Author(s)	押野, 智博
Description	配架番号：2981
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第16888号
Issue Date	2026-03-25
DOI	https://doi.org/10.14943/doctoral.k16888
Doc URL	https://hdl.handle.net/2115/99733
Type	doctoral thesis
File Information	OSHINO_Tomohiro.pdf, 全文



学 位 論 文

造影超音波検査による早期乳癌の腋窩リンパ節転移診断に関する研究：
人工知能併用および非併用での検証 (Diagnosis of Axillary Lymph Node
Metastasis in Early Breast Cancer Using Contrast-Enhanced Ultrasound:
Independent Evaluation of Approaches With and Without Artificial Intelligence
Analysis)

2026年3月

北海道大学

押 野 智 博

学 位 論 文

造影超音波検査による早期乳癌の腋窩リンパ節転移診断に関する研究：
人工知能併用および非併用での検証 (Diagnosis of Axillary Lymph Node
Metastasis in Early Breast Cancer Using Contrast-Enhanced Ultrasound:
Independent Evaluation of Approaches With and Without Artificial Intelligence
Analysis)

2026年3月

北海道大学

押 野 智 博

目 次

発表論文目録および学会発表目録	1 頁
要旨	3 頁
略語表	6 頁

第 1 章

造影超音波検査を用いた腋窩リンパ節転移診断能の評価：人工知能を非併用
-Discrepancy の検証-

緒言	7 頁
方法	9 頁
結果	13 頁
考察	22 頁

第 2 章

造影超音波検査を用いた腋窩リンパ節転移診断能の評価：人工知能を併用

緒言	24 頁
方法	27 頁
結果	39 頁
考察	47 頁

全体の結論	50 頁
謝辞	52 頁
利益相反	53 頁
引用文献	54 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. Oshino T, Enda K, Shimizu H, Sato M, Nishida M, Kato F, Oda Y, Hosoda M, Kudo K, Iwasaki N, Tanaka S, Takahashi M

Artificial intelligence can extract important features for diagnosing axillary lymph node metastasis in early breast cancer using contrast-enhanced ultrasonography

Scientific Reports 2025; 15(1):5648. IF 3.9

2. Oshino T, Shimizu H, Sato M, Nishida M, Horie T, Tsuneta S, Kato F, Hosoda M, Yokota I, Kudo K, Takahashi M

Impact of breast tumor size discrepancy between contrast-enhanced and conventional ultrasonography on axillary node metastasis: a retrospective cohort study

BMC Cancer 2025; 25(1):1718. IF 3.4

本研究の一部は以下の学会に発表した。

1. 押野 智博, 竹下 卓志, 守谷 結美, 加藤 扶美, 高橋 将人

深層学習を用いた造影超音波による原発性乳癌の腋窩リンパ節
転移診断能の検討

第30回日本乳癌学会学術総会 2022年6月30日～7月2日 横浜

2. 押野 智博, 遠田 建, 清水 寛和, 荒町 優香里, 小金澤 千夏, 羽田 光輝, 守谷 結美, 細田 充主, 加藤 扶美, 工藤 與亮, 岩崎 倫政, 高橋 将人

人工知能を用いた造影超音波検査による、原発性乳癌腋窩リンパ節転移の診断能の
検証

第61回日本癌治療学会学術集会 2023年10月19日～21日 横浜

3. 押野 智博, 遠田 建, 清水 寛和, 荒町 優香里, 小金澤 千夏, 李 東, 羽田 光輝, 守谷 結美, 加藤 扶美, 細田 充主, 工藤 與亮, 岩崎 倫政, 高橋 将人

早期乳癌における造影超音波検査での腋窩リンパ節転移診断に重要な所見のラジ
オミクスによる抽出

第124回日本外科学会定期学術集会 2024年04月 名古屋

4. 押野 智博、荒町 優香里、小金澤 千夏、守谷 結美、羽田 光輝、細田 充主、高橋 將人

乳癌の腋窩リンパ節微小転移は予後に影響し、造影超音波検査の原発巣の腫瘍径により予測しうる

第32回日本乳癌学会学術総会 2024年7月11日(木)～13日(土) 仙台

5. Oshino T, Enda K, Shimizu H, Sato M, Nishida M, Kato F, Oda Y, Hosoda M, Kudo K, Iwasaki N, Tanaka S, Takahashi M

Artificial intelligence can extract important features for diagnosing axillary lymph node metastasis in early breast cancer using contrast-enhanced ultrasonography

47th San Antonio Breast Cancer Symposium®, December 10-13, 2024, at the Henry B. Gonzalez Convention Center, San Antonio, Texas.

6. Oshino T, Shimizu H, Sato M, Nishida M, Horie T, Tsuneta S, Kato F, Hosoda M, Yokota I, Kudo K, Takahashi M

Impact of breast tumor size discrepancy between contrast-enhanced and conventional ultrasonography on axillary lymph node metastasis

48th San Antonio Breast Cancer Symposium®, December 9-12, 2025, at the Henry B. Gonzalez Convention Center, San Antonio, Texas.

要旨

【背景と目的】

早期乳癌における腋窩リンパ節転移は乳癌の治療方針や予後に影響する。腋窩リンパ節への手術が施行されれば腋窩リンパ節転移の有無も病理診断により最終診断されるものの、腋窩リンパ節への手術には合併症もあるため、腋窩手術は縮小傾向である。ここで画像診断が重要となり、不適切な画像診断は、不適切な腋窩リンパ節への手術や術後治療につながる。これらから、腋窩手術とその後の病理診断を行わずに腋窩リンパ節転移の状態を予測する正確で非侵襲的な手段、とりわけ画像診断が必要と考えられる。

腋窩リンパ節に対する画像診断法の中で、超音波検査 (US)は安価で簡便であり、放射線被曝や侵襲も伴わないため、世界中で利用されている。しかし非造影 US による腋窩リンパ節転移診断能は十分ではなく、さらに統一されたコンセンサスも得られておらず、診断能の報告もばらつきが大きい。一方で造影超音波検査 (造影 US)は乳癌において保険診療として利用されており、腋窩リンパ節転移の診断能は非造影 US よりも高いとする報告が散見され、非造影 US の利点を保ちながら高い診断能を有することが期待される。しかし、診断基準のコンセンサスは非造影 US 以上に少なく、診断能も向上が求められる。

今回、造影 US による腋窩リンパ節転移診断の向上・診断の統一化のために、2つのアプローチによる検証を行った。第1章は、背景として乳癌原発巣の非造影 US 画像を用いた腋窩リンパ節転移の予測が近年報告されていること、造影 US における非造影 US に比べた病変範囲の拡大 (以下 Discrepancy)は豊富な血管構築を反映し転移能を示す可能性があることに注目し、新たな診断基準の探索として Discrepancy が腋窩リンパ節転移の予測因子になるかを評価することを目的とした研究である。第2章は人工知能の医療への導入に注目した。画像診断では人工知能が広く用いられており、深層学習モデルは画像データを直接入力して高い精度で人工知能の予測を得ることができる。実際に、乳癌原発巣の非造影 US 画像を利用して腋窩リンパ節転移を深層学習により予測する試みなども行われているものの、造影 US 画像を深層学習に用いた研究は報告されていなかった。一方で機械学習モデルは特徴量による最適な予測を提供しながら特徴量の間の複雑な相互関係を明らかにすることができ、医療画像の特徴量を抽出することで医療画像にも活用できる。特に代表的なツリーベースモデルである特徴選択を用いた Light-gradient boosting machine (以下 LightGBM)は表形式の大規模データセットと高次元特徴から学習できる重み付き決定木モデルで、高精度な分類と特徴抽出を行うことができる。しかし、これらの機械学習モデルを用いて造影 US 画像を解析した報告はなかった。これらを背景に、人工知能の導入による造影 US の腋窩リンパ節転移診断能の評価と、診断のための重要な特徴量の抽出を目的とした研究を第2章として行った。

第1章

【対象と方法】

2013年～2021年に当院で早期乳癌 (Stage 0-3)の手術を受けた症例のうち、非浸潤性乳管癌の症例、術前薬物療法を施行した症例、術前の造影USが未施行な症例、US画像での腫瘍径評価が困難な症例、非造影USで腫瘍径が20mm未満または50mmを超える症例を除いた259例を対象とした。①造影USと非造影USの腫瘍径差 (Discrepancy)が4mm以上のDISCR群94例と、②それ以外のnon-DISCR群165例に分けて腋窩リンパ節転移の有無ならびに予後を評価した。

【結果】

DISCR群はnon-DISCR群に比べて腋窩リンパ節転移陽性の割合が有意に高く ($P<0.001$), 多変量解析でも独立した予測因子であった (オッズ比 5.838, 95% 信頼区間 2.408-14.155)。一方造影USの腫瘍径は有意な予測因子とならなかった (オッズ比, 0.984; 95% 信頼区間, 0.939-1.030)。また、Discrepancyは定量的にも腋窩リンパ節転移の独立した予測因子であった (オッズ比: 1.106, 95% 信頼区間, 1.022-1.197)。

更に5年無再発生存期間は、DISCR群 0.750 (95% 信頼区間 0.632-0.868)、non-DISCR群 0.924 (95% 信頼区間 0.870-0.978)と、DISCR群で有意に予後が悪かった。

【考察】

造影USによるDiscrepancyが腋窩リンパ節転移を予測し得るのであれば、様々な臨床活用が想定される。特に、腋窩リンパ節の画像所見や他のモダリティとも組み合わせ、センチネルリンパ節省略の症例を適切に評価することに活用が期待される。一方、造影USの主観性・専門性・時間的コストなどの課題も残り、これらの解決方法も同時に検討していく必要がある。

第2章

【対象と方法】

2013年～2021年に当院で早期乳癌 (Stage 0-3)の手術を受けた患者のうち、造影USで腋窩リンパ節を評価しなかった症例、腋窩手術が行われなかった症例を除いた788例を解析対象とした。本研究は2つのステップによる人工知能解析である。①深層学習モデルを用いて非造影USと造影US画像を重ね合わせた画像を入力し、画像による腋窩リンパ節転移の予測共変量 (iP:0-1)を出力した。②機械学習モデルを用いて、画像の所見を医師が読影し表形式データ化した表形式データと深層学習の予測値 (iP)を組み合わせてLightGBMに学習させ、最終的な予測値を出力した。さらに、それぞれの特徴量の重要度も評価した。

【結果】

①深層学習モデルでは、EfficientNet B8がArea under the receiver operating characteristic curve (AUC) 0.687であり他のモデルと比べて最も高く、機械学習への入力データにも利用した。

②機械学習モデルでは、AUC は造影 US を追加した場合に 0.93 (0.88-0.98)と、非造影 US 単独の 0.76 (0.68-0.85)よりも高かった ($p<0.001$, DeLong test)。

また、重要度が高い特徴は (1)不均一な造影; 3.24×10^2 、(2)びまん性の皮質肥厚 $>3\text{mm}$; 3.16×10^2 、(3)偏心性の皮質肥厚 $>3\text{mm}$; 2.58×10^2 であった。また、深層学習の予測 iP は 5.14×10^1 と 6 番目の高い重要度であり、機械学習の診断に寄与し得た。

診断を簡略化するためにこの 3 つの特徴のみを用いた腋窩リンパ節転移の診断能は、陽性所見が 1 個の場合に精度 0.90、感度 0.68、特異度 0.95、2 個の場合に精度 0.86、感度 0.43、特異度 0.98 となった。

【考察】

①深層学習モデルによる診断は診断能向上の余地がある。診断能向上のために造影 US の動画を利用する、症例数を増やすなど、様々な要因から改善が期待できる。

②機械学習モデルは高い診断能を得られたものの、医師の所見評価が加わっておりエンドツーエンド (end-to-end)の研究では無く、主観性の改善が今後求められる。

【全体の結論】

非造影 US と造影 US との乳癌原発巣の大きさの相違 (Discrepancy)は、腋窩リンパ節転移の独立した予測因子であり、無再発生存期間などの予後とも関連していた。深層学習モデルを用いた造影 US 画像による早期乳癌の腋窩リンパ節転移予測は、機械学習診断の補助になり得るものの、改善の余地がある。機械学習を用いた解析は、造影 US の追加により高い診断能を得ることができた。さらに”不均一な造影効果”や”びまん性の皮質肥厚”、“偏心性の皮質肥厚”などの重要な特徴量を抽出することができた。

乳癌腋窩リンパ節転移診断の画像診断向上のために、今後も造影 US の追加の可能性を、人工知能併用の可能性と併せて模索していきたい。

略語表

本文中および図中で使用した略語は以下の通りである。

ACGAN	auxiliary classifier generative adversarial nets
AUC	area under the receiver operating characteristic curve
BI-RADS	breast imaging reporting and data system
DICOM	digital imaging and communications in medicine
DISCR	group of patients with discrepancy $\geq 4\text{mm}$
ER	estrogen receptor
GAN	generative adversarial network
HER2	human epidermal growth factor receptor 2
LightGBM	light gradient boosting machine
MRI	magnetic resonance imaging
PACS	picture archiving and communication system
PgR	progesterone receptor
PMRT	postmastectomy radiation therapy
ROC	receiver operating characteristic curve
US	ultrasonography

第1章—造影超音波検査を用いた腋窩リンパ節転移診断能の評価—

人工知能解析を非併用 -Discrepancy の検証- —

背景

乳癌の診療における問題点

乳癌は女性に最も多い悪性腫瘍であり、世界的には死亡率は2番目に高い (Siegel RL et al, 2024)。さらに、早期乳癌における腋窩リンパ節転移は乳癌の予後に影響する (Soerjomataram I et al, 2008; Chang JM et al, 2020)。腋窩リンパ節への手術が施行されれば腋窩リンパ節転移の有無も病理診断により最終診断されるものの、腋窩リンパ節への手術には合併症もあるため (Langer I et al, 2007)、腋窩手術は縮小傾向である。具体的にはN0症例へ腋窩リンパ節郭清を可能な限り省略してセンチネルリンパ節生検で代用することはもちろん、近年ではT1 (Gentilini OD et al, 2023) やT2 (Reimer T et al, 2025)かつN0症例の手術先行におけるセンチネルリンパ節生検の省略や、N+症例の術前化学療法症例後の腋窩郭清省略 (Kang YJ et al, 2017)なども標準化されつつあり、これらの診断には画像診断が重要となっている。すなわち不適切な画像診断は、不適切な腋窩リンパ節への手術や術後の治療につながる。したがって、腋窩手術を行わずに腋窩リンパ節転移の状態を予測する正確で非侵襲的な手段が必要である。

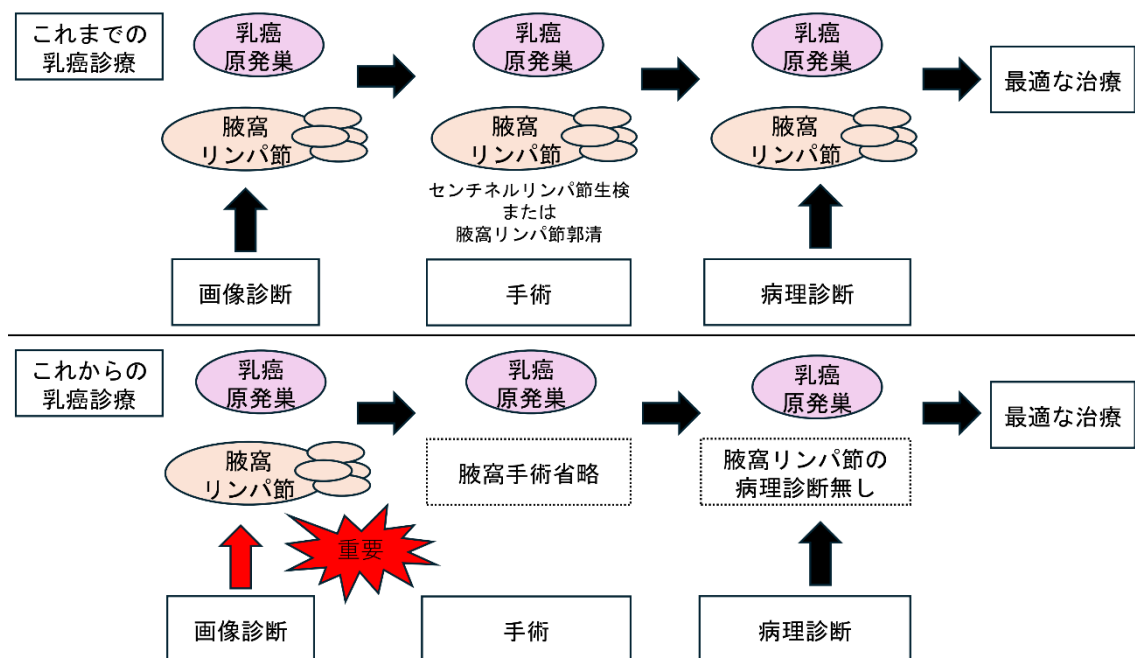


図1. 腋窩リンパ節造影の画像診断の重要性
腋窩リンパ節の画像診断の課題

腋窩リンパ節に対して利用可能な様々な画像診断法の中で、超音波検査 (US)は安価で簡便であり、放射線被曝や侵襲を伴わないため、世界中で利用されている (Marino MA et al, 2020)。非造影 US の診断能は、感度 49-87%、特異度 56-97%、精度 71-84%と報告されている (Jin L et al, 2022; Sato K et al, 2004; Mills P et al, 2010)。びまん性皮質肥厚 (Garcia-Ortega MJ et al, 2011)など腋窩リンパ節転移の評価には様々な指標があるものの、統一されたコンセンサスは得られておらず、診断能の報告もばらつきが大きい。

一方造影超音波検査 (造影 US)は乳癌において保険診療として利用されており、腋窩リンパ節転移の診断能は感度 61-100%、特異度 82-100%、精度 80-92%と、非造影 US よりも高いとする報告が散見され (Jin L et al, 2022; Du LW et al, 2021; Liu YB et al, 2021; Cox K et al, 2013)、非造影 US の利点を残しながらより高い診断能を得られる可能性がある。そのため、今回我々は造影 US による腋窩リンパ節転移の診断に注目した。しかし診断能は未だ十分ではなく、加えて非造影 US よりも多くの診断指標 (Ouyang Q et al, 2010)があり非造影 US 以上に標準的な撮像法・読影法が確立されていないため、普及率は低い。造影 US を普遍的かつ効果的なものとするためには、より簡便で診断能の高い指標が求められる。

乳癌原発巣の腋窩リンパ節診断への利用と、造影 US による腫瘍径の増大 (Discrepancy)について

乳癌原発巣の非造影 US 画像を用いて腋窩リンパ節転移を予測することも近年報告されている (Zheng X et al, 2020)。さらに造影 US は造影剤を用いて血管と周辺組織のコントラストを増強することで、病変部の微小循環灌流に関する情報や、非造影 US では検出できない血管の数、太さ、形状、分布などの特徴を提供することができ的可能性がある (Cantisani V et al, 2021; Rafailidis V et al, 2020)。乳腺腫瘍では、造影 US による腫瘍径の拡大は良性よりも悪性腫瘍で起こりやすいと報告されており (Jiang YX et al, 2007)、この現象は悪性腫瘍との関連が示唆されている。さらに、悪性病変の浸潤前面 (Invasive front)の血管リモデリングが転移に重要な役割を果たすことが示されている (Sanegre S et al, 2021; Bhargava A et al, 2023)。これらから、造影 US と非造影 US の間の腫瘍径の不一致 (Discrepancy)は、病変の転移能を反映している可能性があるかと仮定した。

本研究の目的

本研究では、早期乳癌において造影 US を追加した場合に観察される腫瘍サイズの増大と腋窩リンパ節転移との関連を調べることを目的とした。

方法

倫理審査・指針・説明同意文書

本研究は、既存の情報のみを利用した単施設後方視研究である。自 020-0392
「造影超音波による乳癌腋窩リンパ節転移診断能の検討」として、当院の倫理審査委員会の承認を得ている。

事前に取得した包括同意取得文書の「乳がんの適切な治療法の開発に関する遺伝子解析研究」または「北海道大学病院 乳腺外科：診療目的で採取された血液・組織や、研究目的に提供いただく血液などの研究用保管とその利用について」+オプトアウトにより、書面によるインフォームド・コンセントの要件は免除されている。

これらの承認のもとで「人を対象とした医学系研究に関する倫理指針」に従い解析している。

症例集積 (図 2)

対象者

2013 年～2021 年に当院で早期乳癌 (Stage 0-3)の手術を受けた 971 例

除外基準

- ①非浸潤性乳管癌 (Ductal carcinoma in situ, Stage 0)の症例
- ②術前薬物療法を施行した症例
- ③術前の造影 US が未施行または US 画像での腫瘍径評価が困難な症例
- ④非造影 US で腫瘍径が 20 mm 未満または 50 mm 超える症例

グループ分け

- ①DISCR : 造影 US と非造影 US の腫瘍径差 (Discrepancy)が 4mm 以上, すなわち (造影 US 腫瘍径) – (非造影 US 腫瘍径) ≥ 4.0
- ②non-DISCR : DISCR に当てはまらないもの
(4.0 mm の閾値は過去の報告より引用した (Jiang YX et al, 2007)).

症例情報の収集

症例情報は診療録から後ろ向きに収集し、以下の項目を含めた。

臨床因子 : 年齢、手術日、治療内容 (手術、術前・術後の薬物療法、放射線療法)、最終受診日、局所再発・遠隔再発の有無と再発日

画像検査 : 超音波検査、造影超音波検査の所見

病理学的因子 : 腫瘍の組織型、腫瘍径、脈管侵襲、グレード、エストロゲン受容体 (estrogen receptor : ER)、プロゲステロン受容体 (progesterone receptor : PgR)、ヒト上皮成長因子受容体 2 (human epithelial growth factor 2 : HER2)、Ki-67、腋窩リンパ節転移の数と大きさ

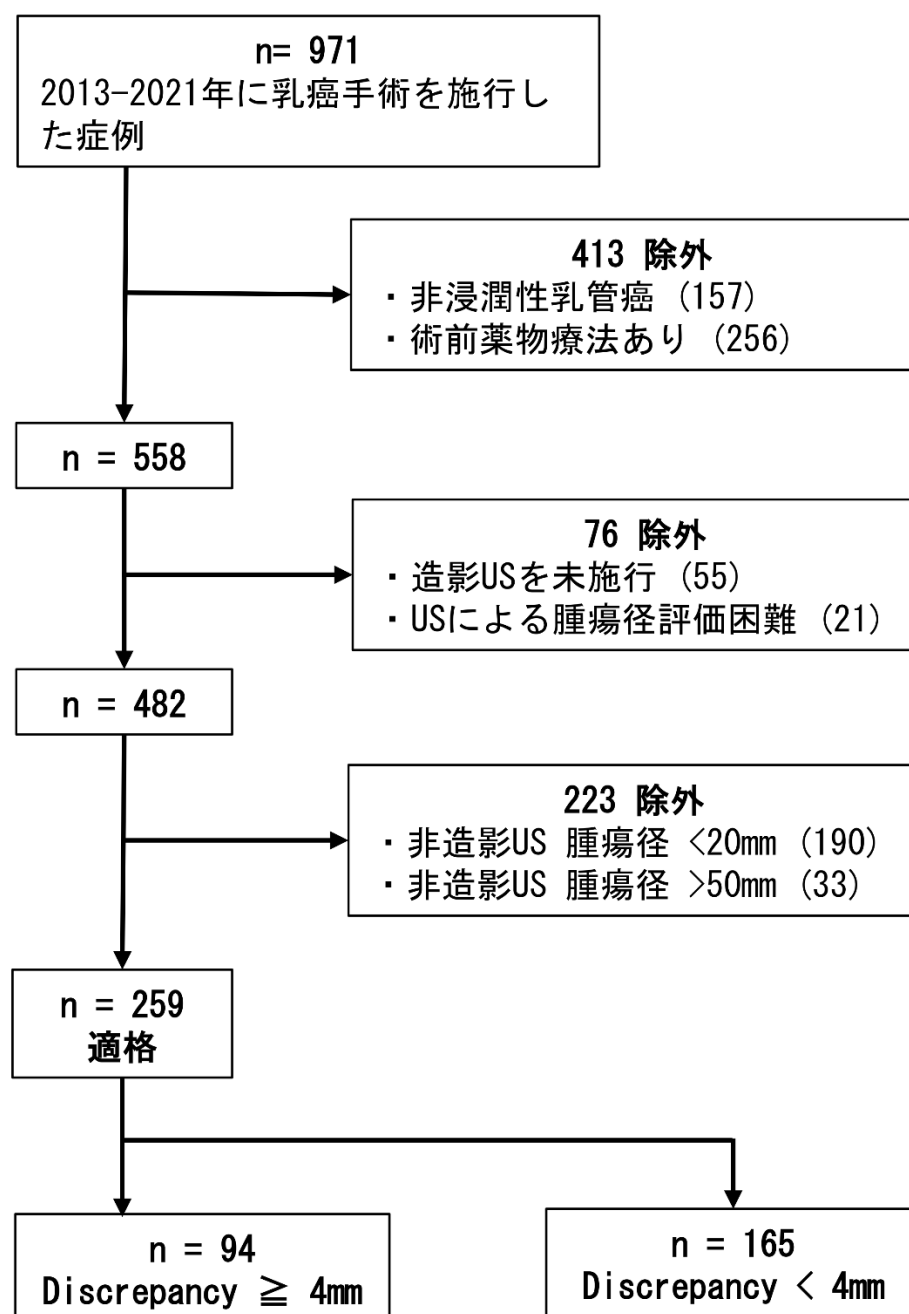


図2 第1章の研究デザイン

造影超音波検査

造影超音波検査は Sonazoid® (GE HealthCare Pharma 社製)0.015 mL/kg/body を静脈注射し、超音波の1つの面で固定して造影剤の流入と流出を評価するものである。当院では診療の一環として行われ、乳癌の診断となった症例に原則施行された。1

度の検査で2回の造影剤注射が行われ、乳癌原発巣における造影剤の流入出を最大径の面での評価を必ず施行した。

腫瘍径の評価には、造影剤の集積が最大となる瞬間 (概ね造影剤静注後 60 秒以内)が利用された。

検査には乳腺超音波の経験を5年以上有する超音波技師が関与し、読影は放射線科医によって行われた。

腫瘍径の測定

乳癌の腫瘍は様々な形状を呈するため、形状に応じて以下のような計測方法で計測した。

●非造影 US (Jiang YX et al, 2007)

- (1) 円形の境界明瞭な病変は、腫瘍の低エコー領域を測定した
- (2) 微小隆起、角化、棘突起を有する病変は、1つの微小隆起、角化、棘突起の先端から対向する棘突起の先端までの距離を測定した。

●造影 US (Fujisawa T et al, 2013)

最大病変径は最大増強画像から手動で測定した。この際に、造影剤注入後の低エコー領域が非造影 US と比較して増強されることを造影ありと定義した。

非造影 US 上の腫瘍領域と、これに隣接するすべての増強成分を対象とした。

これらを踏まえて、図3のように Discrepancy を測定した。

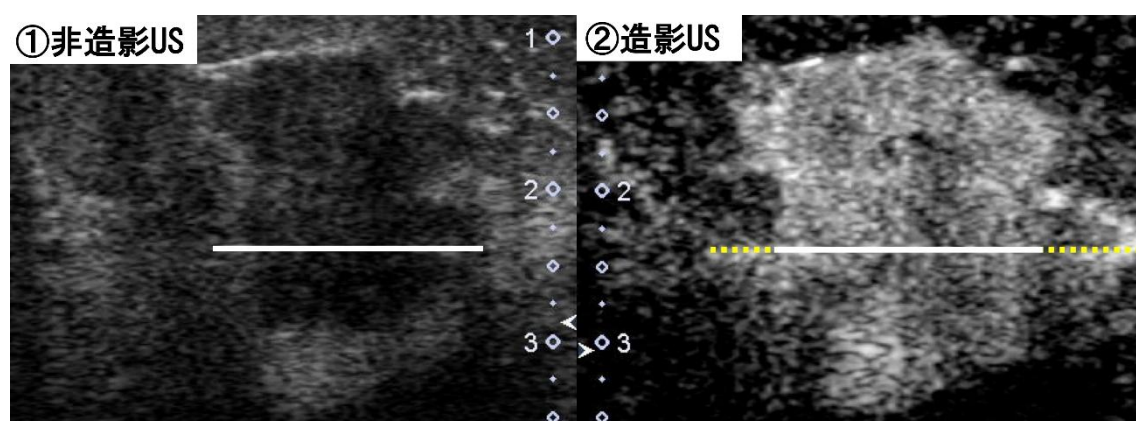


図3 DISCR 群の代表的な超音波画像

65 歳女性、腋窩リンパ節転移を有する乳癌の診断となった。

- ① 非造影 US の腫瘍径
- ② 造影 US の腫瘍径

白の実線は非造影 US による測定範囲で (21 mm)、黄色の点線が造影 US における腫瘍径の差 (Discrepancy)となる (6 mm). 腫瘍の増大が 4 mm 以上であり、この症例は DISCR 群に分類された。

評価した非造影 US・造影 US の所見

腫瘍径以外にも腋窩リンパ節転移の診断に寄与し得る所見を探索するため、原発巣の所見を以下の項目で評価した。

非造影 US

・ Breast imaging reporting and data system (BI-RADS) カテゴリー (Carl J DO et al, 2013)

Category 1: 所見なし、Category 2: 良性、Category 3: おそらく良性 (経過観察推奨、悪性 2%以下)、Category 4: 悪性疑い (生検推奨、悪性 2%-95%)、Category 5: 悪性を強く疑う (生検推奨、悪性 95%以上)

造影 US

・ 造影超音波による原発巣の所見 (カッコ内は陽性所見) (Xiao X et al, 2014)。

A1)造影時間 (周囲乳腺よりも早期に造影)、A2)造影強度 (周囲乳腺よりも強く造影; 明確な基準は無いが明らかに強い場合)、A3)造影方向 (辺縁から求心性の造影)、A4)内部の均一性 (内部不均一な造影)、A5)造影病変の境界 (境界不明瞭)、A6)病変の形状 (不整型)、A7)リング状造影効果 (リング状造影効果あり)、A8)病変の範囲 (長径・短径共に拡大あり)、A9)造影欠損 (造影欠損あり)

解析手法

単変量解析は、各所見および腋窩リンパ節転移の状況を DISCR 群と非 DISCR 群とで比較するために用いた

多変量解析は、Discrepancy・US 所見・臨床病理学的因子などから腋窩リンパ節転移を予測する因子を測定するために用いた。

受信者動作特性曲線 (Receiver operating characteristic curve; ROC 曲線)とその曲線下面積 (Area under the receiver operating characteristic curve; AUC)は本研究における 4mm の閾値の検証に利用された。

評価項目

主要評価項目：

- ・ 腋窩リンパ節転移の病理学的評価

副次的評価項目：

- ・ 60 ヶ月の無再発生存期間 (手術から、最初の所属リンパ節または遠隔再発までの期間)
- ・ 60 か月の全生存期間 (手術から、あらゆる原因による死亡までの期間)

統計解析

カテゴリー変数はカイ二乗検定を、連続変数は t 検定を用いて分析した。p 値はボンフェローニ補正により多重比較の調整をした。多変量解析にはロジスティック回帰分析を用いた。生存率の評価は Kaplan-Meier 法を用いた。すべての統計解析は Bell Curve for Excel version 4.02 または EZR version 1.68 を用い、有意水準は 0.05 とした。

結果

患者背景

対象者 971 例のうち除外は以下の通りで、結果 259 例を解析対象とした。

- ①非浸潤性乳管癌 (DCIS)の症例; 157 例
- ②術前薬物療法を施行した症例; 256 例
- ③術前の造影 US が未施行または US 画像での腫瘍径評価が困難; 76 例
- ④非造影 US で腫瘍径が 20 mm 未満または 50 mm 超える症例
 - ・ 20mm 未満; 190 例 (うち、リンパ節転移陽性 30 例)
 - ・ 50mm 超え; 33 例 (うち、リンパ節転移陽性 16 例)

以下の表 1 に、全体の患者背景を示す。

グループ分け

DISCR 群は 94 例 (腋窩リンパ節転移 47 例, 50%)、non-DISCR (腋窩リンパ節転移 31 例, 19%)は 165 例となった。DISCR 群と non-DISCR 群において、非造影 US の腫瘍径には有意差が見られず ($p=0.82$, ボンフェローニ補正後 0.99)、造影 US の腫瘍径には有意差が見られた ($p<0.001$, ボンフェローニ補正後 <0.001)。

表 1 患者背景

背景因子	n = 259
年齢 (歳)(中央値 [範囲])	64 [30-90]
病理学的浸潤径 (mm)(中央値 [範囲])	18 [0.0-100.0]
エストロゲン受容体	
陽性	217 (83.8%)
プロゲステロン受容体	
陽性	194 (74.9%)
HER2	
陽性	33 (12.8%)
Ki-67 (%) (中央値 [範囲])	22.6 [0.0-98.5]
組織型	
浸潤性乳管癌	214 (82.6%)
浸潤性小葉癌	23 (8.9%)
その他	22 (8.5%)

表 2. Discrepancy によるグループ分けにおける、非造影 US および造影 US で評価した画像所見・腫瘍径の状況

	DISCR (n = 94)	non-DISCR (n = 165)	Raw <i>p</i> -value	Bonf adjusted <i>p</i> -value
非造影 US 腫瘍径 (mm) (中央値 [範囲])	29 [20-50]	29 [20-49]	0.82	0.99
非造影 US BI-RADS カテゴリー			0.089	0.99
0-3	7 (7.4%)	8 (4.8%)		
4	42 (44.7%)	109 (65.7%)		
5	45 (47.9%)	49 (29.5%)		
造影 US 腫瘍径 (mm) (中央値 [範囲])	39 [25-65]	28 [18-52]	<0.001*	<0.001*
乳癌原発巣の造影 US 所見				
A1: 造影時間; 早期	85 (90.4%)	146 (85.0%)	0.54	0.99
A2: 造影強度; 強い	94 (100%)	158 (95.2%)	0.030*	0.36
A3: 造影方向; 多中心性	84 (89.4%)	145 (87.3%)	0.63	0.99
A4: 造影の均一性; 均一	69 (73.4%)	130 (78.3%)	0.37	0.99
A5: 造影境界; 不明瞭	31 (33.0%)	75 (45.2%)	0.062	0.74
A6: 腫瘍の形状 不整型	31 (33.0%)	77 (46.4%)	0.040*	0.48
A7: リング状造影効果; あり	13 (13.8%)	6 (3.6%)	0.0025*	0.03*
A8: 病変の範囲; 拡大	79 (84.0%)	121 (72.9%)	0.046*	0.55
A9: 造影欠損; あり	52 (55.3%)	75 (45.2%)	0.12	0.99

*統計学的有意差あり

US の所見および Discrepancy と腋窩リンパ節転移との関連性

造影 US の定性的評価

DISCR と有意に関連していた因子は以下の通りである (表 2)。

A2 ($p = 0.030$)、A6 ($p = 0.040$)、A7 ($p < 0.01$)、A8 ($p = 0.046$)。ただし、ボンフェローニ調整後は A7 のみが有意であった ($p = 0.03$)

4mm の閾値の裏付け

Discrepancy を 1mm 刻みに分けて、腋窩リンパ節転移を評価した。

AUC 解析により、本症例群でも閾値 4mm が裏付けられた (感度; 0.74、特異度; 0.60、AUC; 0.677 [95% 信頼区間、0.604-0.749]) (図 4)。

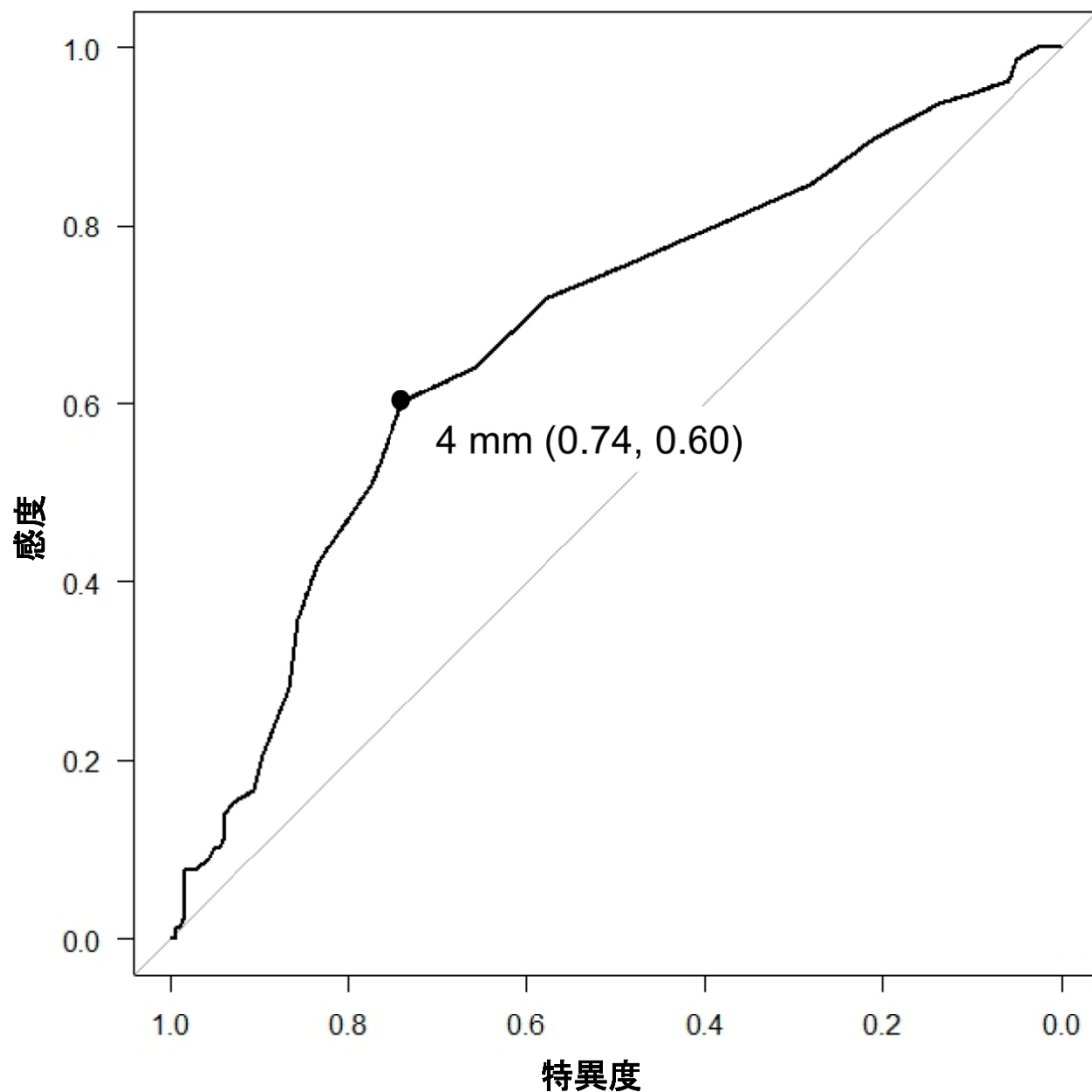


図4 ROC 曲線による Discrepancy の閾値 4mm の裏付け
感度; 0.74、特異度; 0.60、AUC; 0.677 [95% 信頼区間, 0.604–0.749].

Discrepancy と腋窩リンパ節転移の有無の評価

Discrepancy の有無 ($\leq 0\text{mm}$ or $> 0\text{mm}$)および Discrepancy が 4mm 以上か否かで腋窩リンパ節転移の差を評価した (図5)。DISCR 群は、non-DISCR 群に比べて腋窩リンパ節転移の割合が有意に高かった ($p<0.001$)。

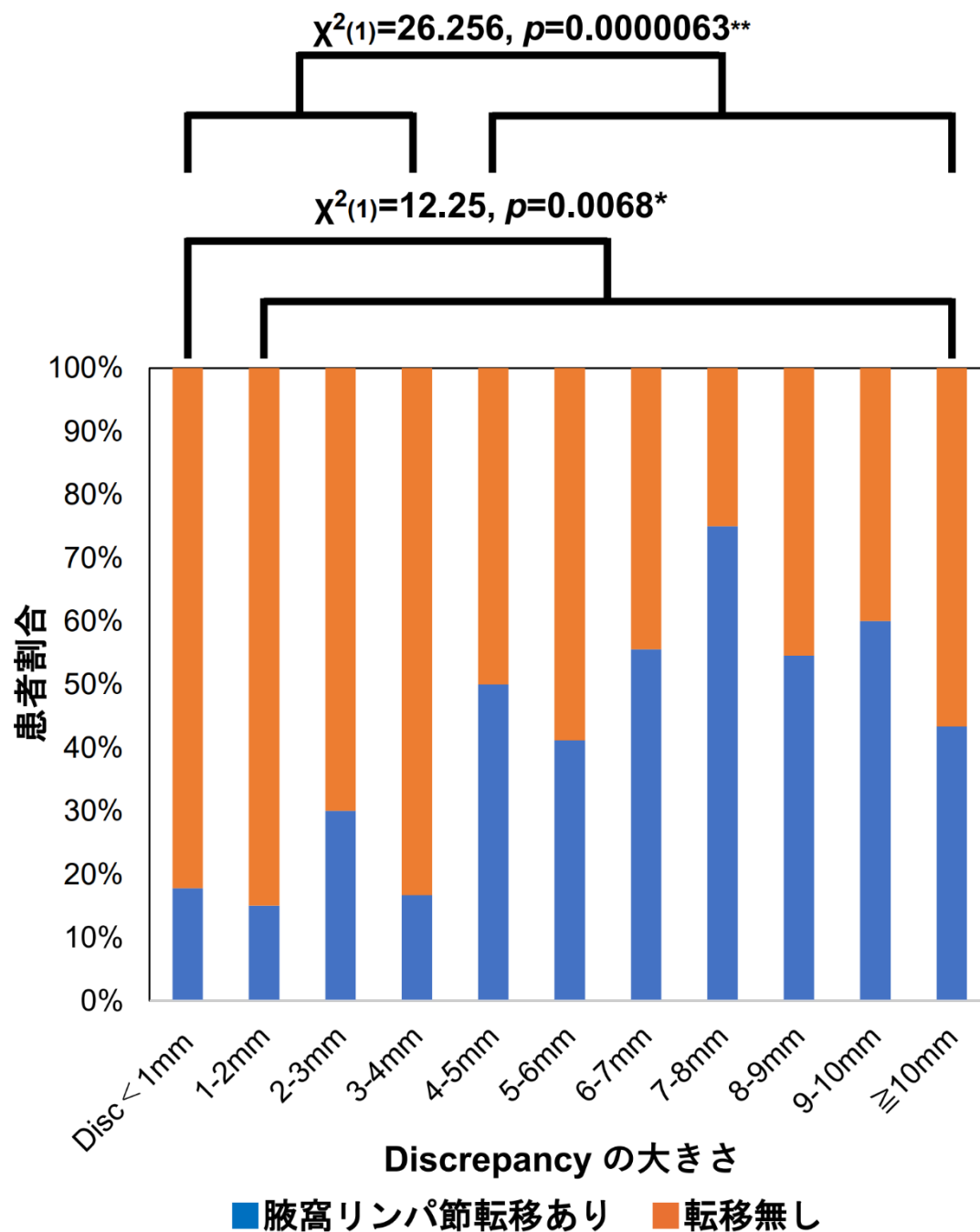


図 5 Discrepancy の大きさごとによる腋窩リンパ節転移の割合
有意差の比較は χ 二乗検定を使用した。

多変量解析

除外症例

16 例は多変量解析から除外された。内訳としては、Ki-67 値の未評価が 9 例、グレードの未評価が 8 例、HER2 の未評価が 1 例、A8 の評価困難が 1 例であった (重複 3 例)。

3 例は腋窩リンパ節転移ありの DISCR 群、3 例は腋窩リンパ節転移ありの non-DICSR 群、10 例は腋窩リンパ節転移無しの non-DICSR 群であり、除外前の割合と概ね同等であるため影響は小さいと考えられた。

多変量解析 (Discrepancy が定性的な場合) (表 3)

Discrepancy を 4mm 以上 (DISCR 群)か否かの定性的な評価にした場合、DISCR 群であることは独立して腋窩リンパ節転移と関連する因子だった (オッズ比: 5.838, 95% 信頼区間, 2.408-14.155), 一方造影 US の腫瘍径は有意な関連する因子とならなかった (オッズ比, 0.984; 95% 信頼区間, 0.939-1.030).

多変量解析 (Discrepancy を定量的にした場合) (表 4)

Discrepancy は定量的にも独立して腋窩リンパ節転移と関連する因子だった (オッズ比: 1.106, 95% 信頼区間, 1.022-1.197).

表 3 腋窩リンパ節転移予測の多変量解析 -Discrepancy が定性的な場合-

	オッズ比	腋窩リンパ節転移	
		95% CI 下限	95% CI 上限
Discrepancy (4.0 mm 以上)*	5.838	2.408	14.155
造影 US 腫瘍径	0.984	0.939	1.030
病理学的浸潤径*	1.073	1.034	1.112
腫瘍グレード	0.772	0.375	1.589
エストロゲン受容体	1.003	0.988	1.018
プロゲステロン受容体	0.994	0.981	1.007
HER2	0.891	0.648	1.227
Ki67	0.997	0.971	1.023
静脈侵襲	1.276	0.511	3.186
リンパ管侵襲*	5.011	1.970	12.748
A2: 造影強度; 強い	0.728	0.038	14.022
A6: 腫瘍の形状; 不整型	0.878	0.359	2.144
A7: リング状造影効果; あり	4.021	0.909	17.785
A8: 病変の範囲; 拡大	1.445	0.453	4.611

年齢, 組織型, BI-RADS カテゴリー, A1, A3, A4, A5, A9 で調整した

*統計学的有意差あり

表 4. 腋窩リンパ節転移予測の多変量解析 -Discrepancy が定量的の場合-

	オッズ比	腋窩リンパ節転移	
		95% CI 下限	95% CI 上限
Discrepancy (mm)*	1.106	1.022	1.197
造影 US 腫瘍径	0.986	0.937	1.039
病理学的浸潤径*	1.068	1.031	1.106
腫瘍グレード	0.779	0.390	1.557
エストロゲン受容体	1.003	0.989	1.018
プロゲステロン受容体	0.994	0.982	1.007
HER2	0.900	0.668	1.213
Ki67	0.998	0.974	1.023
静脈侵襲	1.268	0.527	3.052
リンパ管侵襲*	5.033	2.077	12.196
A2: 造影強度; 強い	0.896	0.048	16.843
A6: 腫瘍の形状; 不整型	1.037	0.437	2.458
A7: リング状造影効果; あり	3.735	0.851	16.391
A8: 病変の範囲; 拡大	1.584	0.508	4.939

年齢, 組織型, BI-RADS カテゴリー, A1, A3, A4, A5, A9 で調整した

*統計学的有意差あり

治療内容と予後

治療内容 (表 5)

DISCR 群の方が腋窩郭清術 ($p = 0.012$)、化学療法 ($p < 0.01$)、乳癌術後放射線療法 (postmastectomy radiation therapy; PMRT) ($p = 0.012$)の施行割合が多かったが、腋窩リンパ節転移を有する割合が高いためと考えられた。

表 5 DISCR 群と non-DISCR 群における治療内容

治療内容	DISCR (n = 94)	non-DISCR (n = 165)	p-value
乳房手術			0.65
乳房全切除術	79 (84.0%)	134 (81.2%)	
乳房部分切除術	15 (16.0%)	31 (18.8%)	
腋窩手術			0.012*
センチネルリンパ節生検	68 (72.3%)	142 (86.1%)	
腋窩郭清	26 (27.7%)	23 (13.9%)	
内分泌療法			0.64

あり	68 (72.3%)	125 (75.8%)	
なし	25 (26.6%)	40 (24.2%)	
不明 ^a	1 (1.1%)	0	
化学療法			0.0036*
あり	42 (44.7%)	45 (27.3%)	
なし	51 (54.3%)	120 (72.7%)	
不明 ^a	1 (1.1%)	0	
抗HER2 療法			0.44
あり	10 (10.6%)	13 (7.9%)	
なし	83 (88.3%)	152 (92.1%)	
不明 ^a	1 (1.1%)	0	
乳房切除後放射線療法 (PMRT)			0.012*
あり	11 (11.7%)	6 (3.6%)	
なし	83 (88.3%)	159 (96.4%)	

*統計学的有意差あり

予後

予後結果は以下の通りとなった。

平均観察期間; 73 か月 (5-125 か月)。

局所再発; DISCR 群 5 例、non-DISCR 群 5 例

遠隔再発; DISCR 群 12 例、non-DISCR 群 7 例

現病死; DISCR 群 2 例、non-DISCR 群 4 例

乳癌以外の死亡; DISCR 群 2 例、non-DISCR 群 3 例 (1 例は遠隔再発後)

5 年無再発生存期間 (RFS) : 有意差あり (図 6)

DISCR 群 0.750 (95% 信頼区間 0.632-0.868)

non-DISCR 群 0.924 (95% 信頼区間 0.870-0.978)

5 年生存期間 (OS) : 有意差無し (図 7)

DISCR 群 0.869 (95% 信頼区間 0.809-0.982)

non-DISCR 群 0.967 (95% 信頼区間 0.930-1.000)

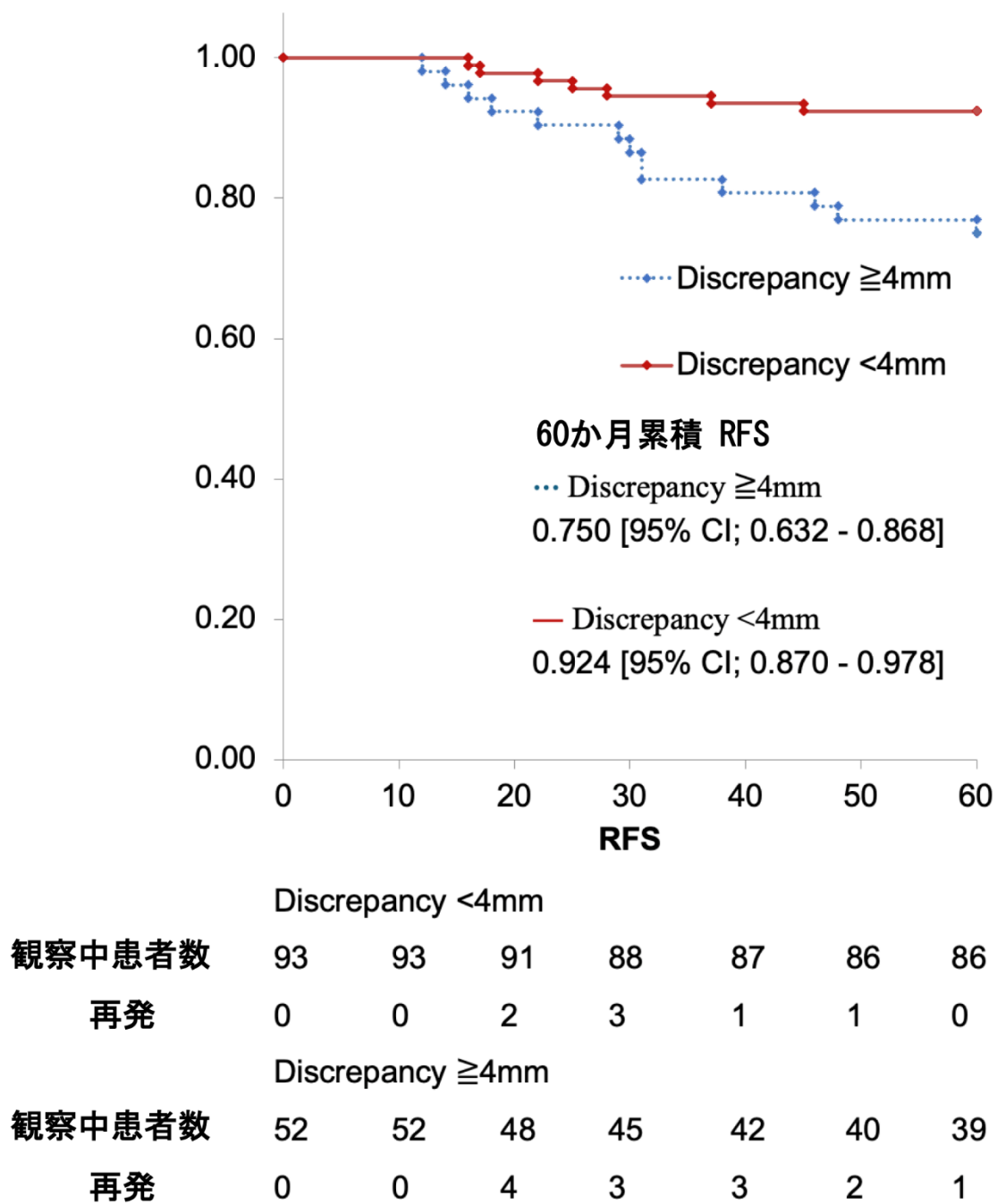


図 6. 60 か月無再発生存期間 recurrence-free survival (RFS)

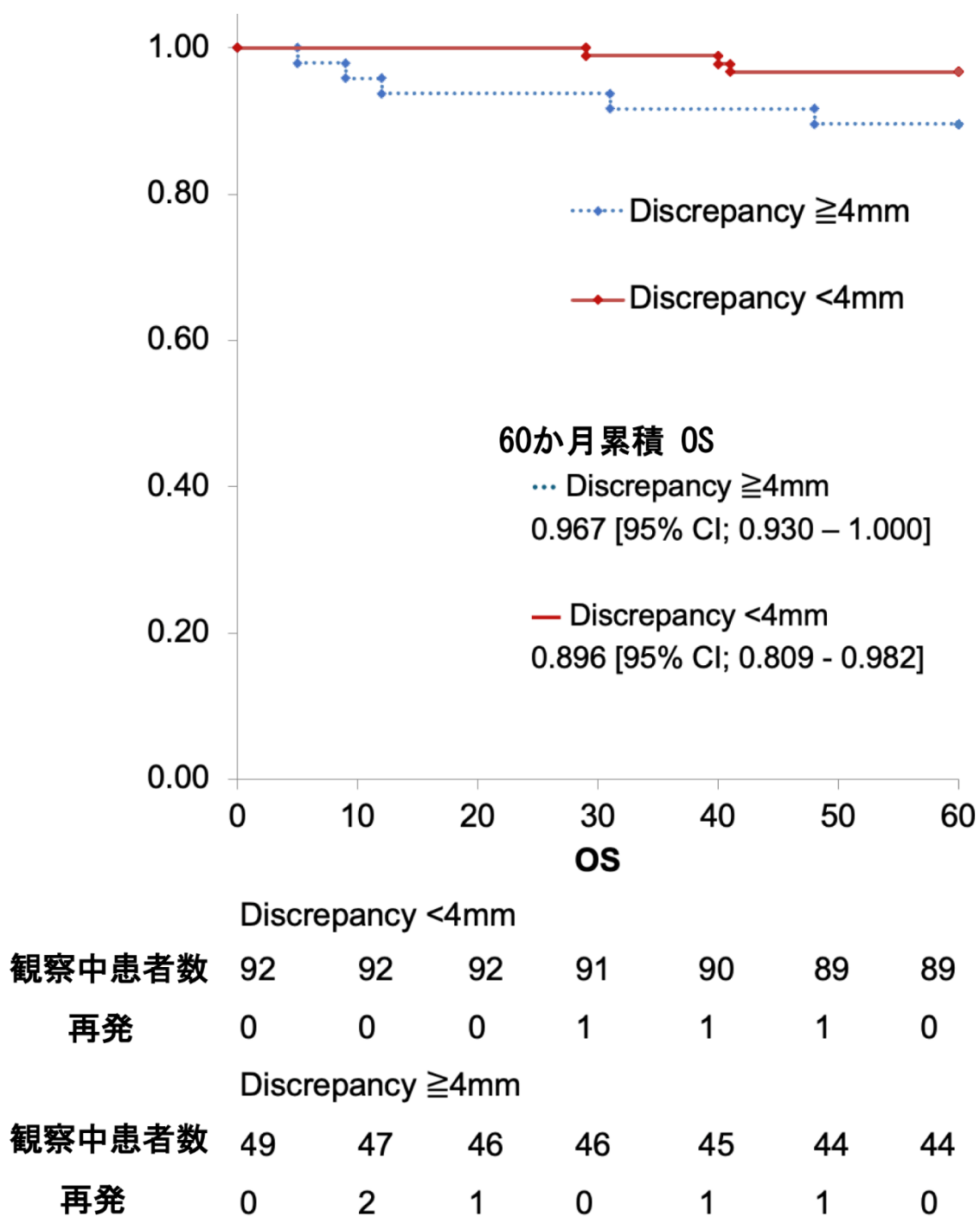


図 7. 60 か月全生存期間 overall survival (OS)

考察

1.本研究で得られた新知見

本研究により、乳癌原発巣における造影 US と非造影 US の大きさの不一致 (Discrepancy) が、腋窩リンパ節転移および予後と関連することが示された。多変量解析の結果、造影 US での腫瘍径は予測因子ではなく、Discrepancy が腋窩リンパ節転移の独立した関連因子であった。

2.先行研究との比較

造影 US の乳癌への利用と Discrepancy の評価について

Discrepancy が腋窩リンパ節転移の予測因子となるならば、臨床にも活用できる可能性がある。造影 US は良性病変と悪性病変の鑑別において造影乳房 magnetic resonance imaging ; MRI (Ito T and Komoike Y, 2024) に匹敵する診断能を有するため、主にこの目的で用いられてきた (Fu Y et al, 2024)。これまでの研究で、乳癌原発巣における造影 US の腫瘍径自体や腫瘍面積と腋窩リンパ節転移との関連性が報告されているものの (Wang Z et al, 2014)、悪性腫瘍に特異的な画像的特徴は同定されていない (Sidhu PS et al, 2018)。本研究では、造影 US による腫瘍径だけでなく、非造影 US と比較した腫瘍径の増大 (Discrepancy) によって腋窩リンパ節転移を予測することで、造影 US の有用性が示された。この造影 US による腫瘍の増大 (Discrepancy) の測定は、腫瘍面積の測定よりも容易であり、臨床での使用に適している。

以前の報告 (Jiang YX et al, 2007) では、Discrepancy の閾値は腫瘍の大きさに依存した; すなわち 17mm 以上の腫瘍では 4mm、17mm 未満では 3mm であった。本研究では 20mm 未満の腫瘍は除外されたが、今後の研究ではより広い腫瘍の大きさを含め、それぞれの腫瘍の大きさに特異的な閾値を採用すべきである。

臨床的な活用方法

造影 US によりセンチネルリンパ節生検の不適切な省略を避けられる可能性がある。昨今ではセンチネルリンパ節生検の省略が盛んに検討されている (Gentilini OD et al, 2023; Reimer T et al, 2025)。しかし、不適切なセンチネルリンパ節生検の省略によりリンパ節転移を見逃すと、局所治療のみならず薬物療法や予後にも影響を及ぼす可能性があり (von Minckwitz G et al, 2017; Johnston SRD et al, 2020; Kalinsky K et al, 2021; Tutt ANJ et al, 2021)、腋窩リンパ節の状態を術前に評価することは重要である。

一方、造影 US は時間を要する検査であり、乳癌手術を受けるすべての患者に実施することは現実的ではない。さらに造影 US の臨床現場での日常的な使用は、費用・導入の有無・検査施行や読影の専門性の高さによって制限される可能性がある。このため臨床においては対症症例を限定し、非造影 US、CT、MRI で腋窩リンパ節転移の所見がなく (cN0)、センチネルリンパ節生検の省略が考慮される乳癌患者にのみへの造影 US の適応が妥当と考えられる。適切な症例における、造影 US 評価に

基づくセンチネルリンパ節生検の省略に関する、前向き研究が必要である。

造影 US による Discrepancy と腋窩リンパ節転移に関する、基礎的な考察

造影 US により腫瘍の増大 (Discrepancy) が観察されることは臨床的に知られている現象であるものの、その基礎となるメカニズム、さらに転移との関連におけるメカニズムはまだ明らかにされていない。

腫瘍生物学では、腫瘍細胞は *invasive front* を形成することによって周囲組織に移動し、浸潤する (Christofori G, 2006)。3 次元微小血管データを用いた異種移植モデルでは、腫瘍の中心部分と比較して *invasive front* での血管新生が豊富であることが示された (Bhargava A et al, 2023)。造影 US を加えると腫瘍が増大するのは、*invasive front* に遺伝的变化 (Novikov NM et al, 2021)、宿主環境への適応 (Fares J et al, 2020)、代謝の違い (Wei Q et al, 2020)、上皮間葉転換・幹細胞様の特徴 (Ribatti D et al, 2020) があることを反映している可能性がある。これらは仮説に過ぎず、血管が豊富な *invasive front* の特徴を前向きに解明するためには、さらなる研究が必要である。

3. 本研究の特徴と問題点

本研究の特徴は、既報よりも多くの症例数で行え、造影 US と非造影 US の腫瘍径の差という評価しやすい項目で、乳癌の腋窩リンパ節転移という臨床的に重要な因子を予測し得ることである。一方で以下の問題点もあり、今後の課題となる。

①50mm を超える腫瘍は測定が困難なため除外し、20mm 未満の腫瘍は腋窩リンパ節転移の頻度が低いため除外した。これは特に選択バイアスとなる可能性がある。

②画像の適格性と画像の特徴が医師の判断により決定されたことが、主観的なバイアスをもたらした可能性がある。造影 US 測定における検査者や装置依存性も観察者バイアスの一因となる可能性がある。

③MRI など他のモダリティを用いた Discrepancy の解析は行わなかった。このことが結果に影響を与えた可能性がある。

4. 第1章の小括と、第2章との関連性

非造影 US と造影 US における乳癌原発巣の大きさの相違 (Discrepancy) は、腋窩リンパ節転移の独立した予測因子であり、予後とも関連していた。生物学的観点からも Discrepancy は転移のメカニズムにおいて重要な役割を果たしている可能性がある。

このように、人工知能の導入無しでも腋窩リンパ節転移を予測する新たな診断基準になり得る所見を得ることができ、更にこれは腋窩リンパ節の所見ではなく原発巣の所見も腋窩リンパ節転移の診断に有用である可能性を示唆した。

しかしながら依然として腋窩リンパ節転移の画像診断は十分な診断能を得ることができない。そのため人工知能の導入によりこの課題を解決できないか検証したため、次章で報告する。

第2章—造影超音波検査を用いた腋窩リンパ節転移診断能の評価—

人工知能を併用—

緒言

乳癌の診療における問題点

第1章で述べたように、早期乳癌における腋窩リンパ節転移は乳癌の予後や治療方針に影響し、腋窩リンパ節転移の診断には画像診断が重要となっている。すなわち不適切な画像診断は、不適切な腋窩リンパ節への手術や術後の治療につながる。したがって、腋窩手術を行わずに腋窩リンパ節転移の状態を予測する正確で非侵襲的な手段が必要である。

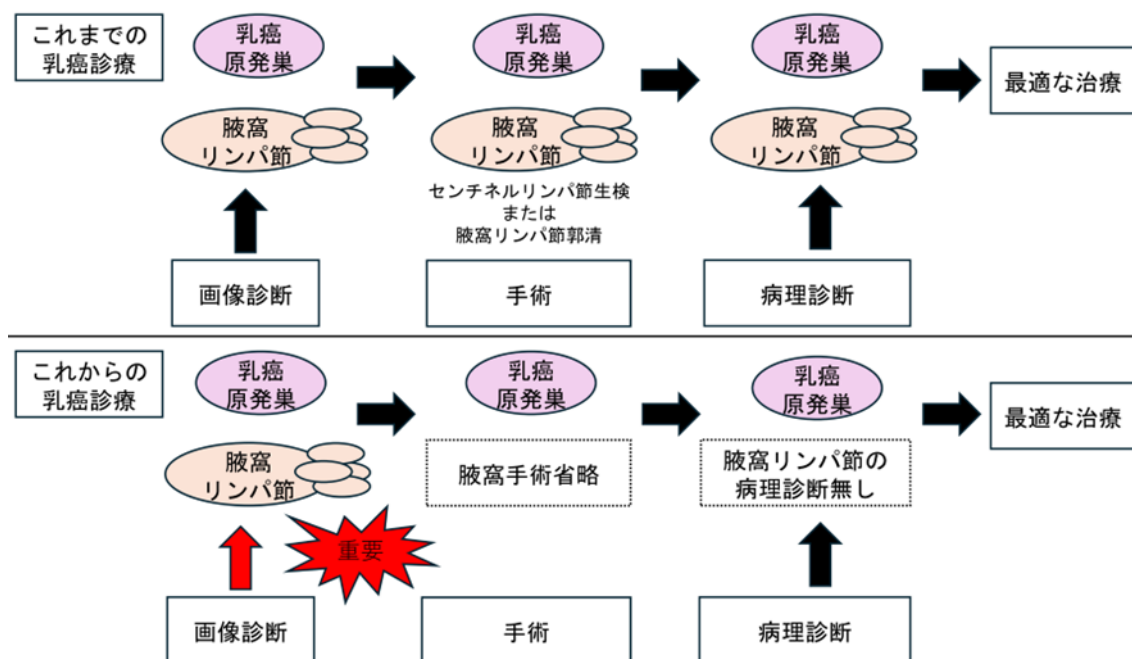


図8. 腋窩リンパ節造影の画像診断の重要性

腋窩リンパ節の画像診断の課題

第1章で述べたように、腋窩リンパ節に対する画像診断法の中で、超音波検査は安価で簡便であり、放射線被曝や侵襲を伴わないため、世界中で利用されている (Marino MA et al, 2020)。非造影 US の診断能は、感度 49-87%、特異度 56-97%、精度 71-84%と報告されている (Jin L et al, 2022; Sato K et al, 2004; Mills P et al, 2010)。しかし、びまん性皮質肥厚 (Garcia-Ortega MJ et al, 2011)など腋窩リンパ節転移の評価には様々な指標があるものの、統一されたコンセンサスは得られておらず、診断能の報告もばらつきが大きい。

造影 US による乳癌腋窩リンパ節転移の診断能は感度 61-100%、特異度 82-100%、精度 80-92%と、非造影 US よりも高い (Jin L et al, 2022; Du LW et al, 2021; Liu YB et al, 2021; Cox K et al, 2013)。非造影 US の利点を残しながらより高い診断能を得られる可能性がある。そのため、今回我々は造影 US による腋窩リンパ節転移の診断に注目した。しかし診断能は未だ十分ではなく、加えて非造影 US よりも多くの診断指標 (Ouyang Q et al, 2010)があり非造影 US 以上に標準的な撮像法・読影法が確立されていないため、普及率は低い。造影 US を普遍的かつ効果的なものとするためには、これらの問題を解決する必要がある。

今回、造影 US による腋窩リンパ節転移診断の向上・診断の統一化のために、人工知能の導入に注目し、人工知能による診断能の評価と診断のための重要な特徴量の抽出を試みた。

人工知能

画像診断では人工知能が広く用いられている (Hosny A et al, 2018)。人工知能の一部に機械学習が、さらに機械学習の一部に深層学習があり (図 9)、それぞれの利点がある。畳み込みニューラルネットワーク (Krizhevsky A et al, 2012)のような深層学習モデルは、画像データを直接的な入出力予測に用いることができる。ただし、特に深層学習モデルによる学習には膨大なデータ数が必要となる。しかし、とりわけ診療上の医療画像データにはその数に限りがあり、不足することが多い。そのため、人工知能の学習データを増量させるための Data augmentation (Buslaev A et al, 2020)という技術が一般的に用いられている。このような技術を応用しながら、乳癌原発巣の非造影 US 画像に基づいて腋窩リンパ節転移を深層学習により予測する試みなども行われている (Zheng X et al, 2020)。しかし、造影 US 画像を深層学習に用いた研究は報告されていなかった。

機械学習モデルは特徴量による最適な予測を提供しながら、特徴量の間の複雑な相互関係を明らかにする学習モデルである。医療画像も含めた画像的特徴は表形式データとすることもでき、表形式データではツリーベースのモデルが深層学習モデルより優れているとされている。代表的なツリーベースモデルである Light-gradient boosting machine (LightGBM) (Ke G, 2017)は表形式の大規模データセットと高次元特徴から学習できる重み付き決定木モデルで、高精度な分類と特徴抽出を行うことができ (Rufo DD et al, 2021; Pan J et al, 2025)、乳房 MRI などにも応用可能である (Vamvakas A et al, 2022)。しかし、これらの機械学習モデルを用い造影 US を解析した報告はなかった。

本研究の目的

本研究の仮説は、造影 US の画像を用いて深層学習モデルによる診断ができること、また造影 US 画像の特徴を機械学習モデルにより診断でき、さらに診断に重要な特徴を抽出できることとした。深層学習または機械学習の LightGBM を用いた乳

癌の腋窩リンパ節転移予測モデルを作成し、造影 US を加えた場合の腋窩リンパ節転移の診断能を評価することを目的とした。また、重要な特徴を抽出し、それらの特徴が臨床的に活用可能かどうかを評価することを目的とした。

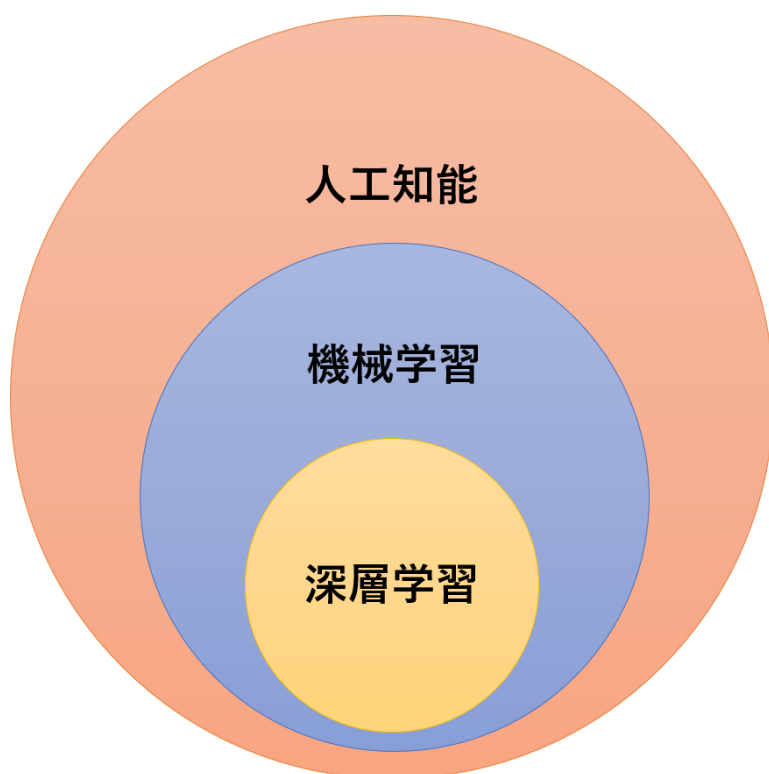


図 9. 人工知能における、機械学習、深層学習の分類

人工知能の一部に機械学習があり、さらにその一部に深層学習がある。

方法

共通事項

倫理審査・指針・説明同意文書

本研究は、既存の情報のみを利用した単施設後方視研究である。自 020-0392 「造影超音波による乳癌腋窩リンパ節転移診断能の検討」として、当院の倫理審査委員会の承認を得ている。

事前に取得した包括同意取得文書の「乳がんの適切な治療法の開発に関する遺伝子解析研究」または「北海道大学病院 乳腺外科：診療目的で採取された血液・組織や、研究目的に提供いただく血液などの研究用保管とその利用について」＋オプトアウトにより、書面によるインフォームド・コンセントの要件は免除されている。

これらの承認のもとで「人を対象とした医学系研究に関する倫理指針」に従い解析している。

症例情報の収集

症例情報は診療録から後ろ向きに収集し、以下の項目を含めた。

臨床因子：年齢、手術日、治療内容（手術、術前・術後の薬物療法、放射線療法）

画像検査：超音波検査、造影超音波検査

病理学的因子：腫瘍の組織型、腫瘍径、脈管侵襲、グレード、エストロゲン受容体 (ER)、プロゲステロン受容体 (PgR)、ヒト上皮成長因子受容体 2 (HER2)、Ki-67、腋窩リンパ節転移の数と大きさ

造影超音波検査

造影超音波検査は Sonazoid® (GE HealthCare Pharma 社) 0.015 mL/kg/body を静脈注射し、超音波の 1 つの面で固定して造影剤の流入と流出を評価するものである。

当院では診療の一環として行われ、乳癌の診断となった症例に原則施行された。1 度の検査で 2 回の造影剤注射が行われ、1 回目は乳癌原発巣が評価され、2 回目は腋窩リンパ節 (図 10) ないしは他に乳癌が疑われる乳腺腫瘍が評価された。腋窩リンパ節が 2 回とも主な評価病変とならなかった場合、腋窩リンパ節の評価は遅延層 (late phase) でのみ行われた。原発巣・腋窩リンパ節どちらも最大径の面で造影剤の流入出を評価した。

腫瘍径の評価や人工知能解析用の画像など、1 枚の瞬間的な静止画が必要な場合には、造影剤の集積が最大となる瞬間 (概ね造影剤静注後 60 秒以内) が利用された。

検査には5年以上乳腺超音波の経験を積んだ超音波技師が関与し、読影は放射線科医によって行われた。

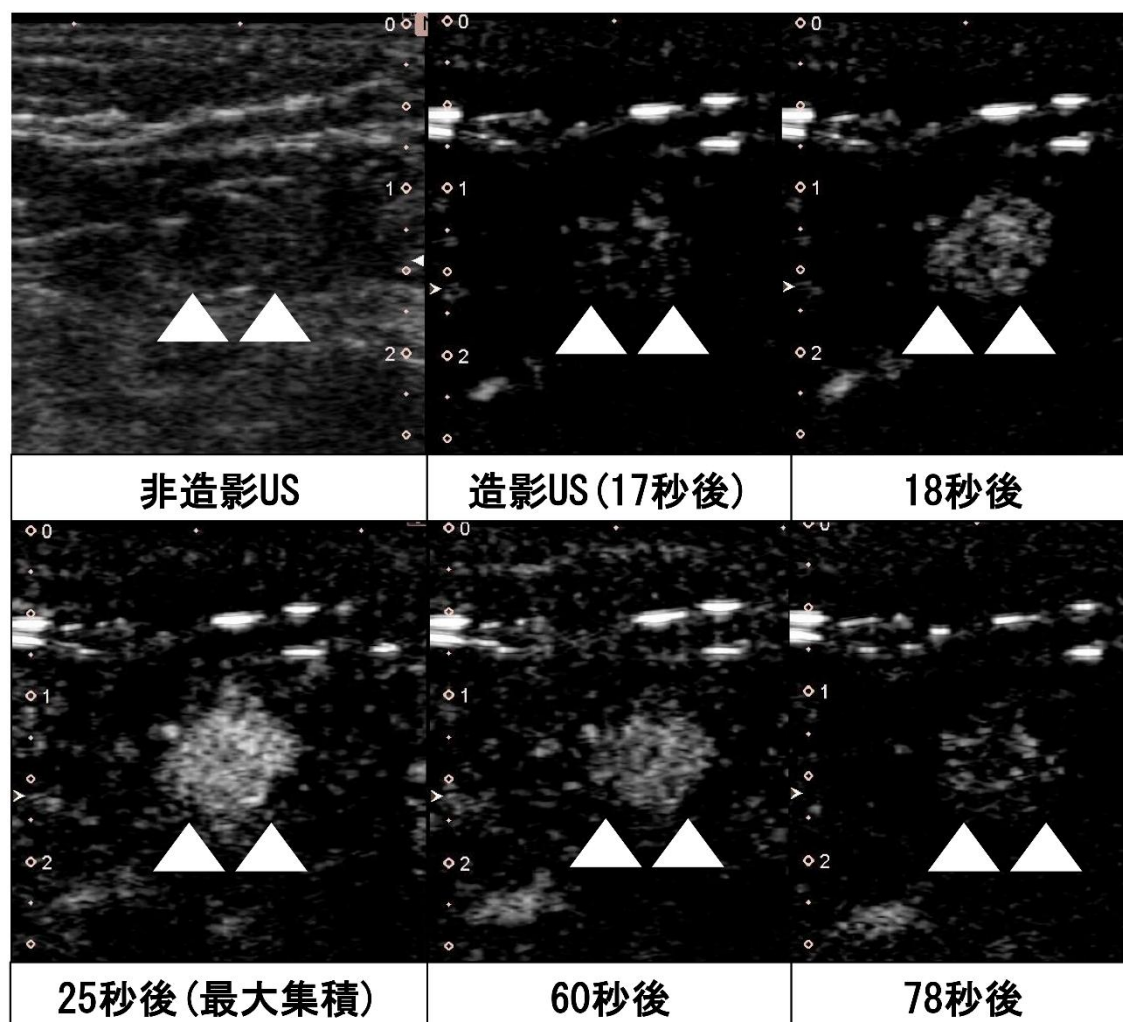


図 10. 造影 US 検査の腋窩リンパ節造影の画像例

US 画像データ取得

各症例の Picture archiving and communication system (PACS) から非造影 US 画像、造影 US 画像を匿名化の上で Digital imaging and communications in medicine (DICOM) データとして入手した。造影 US 画像は、一連の画像の中で最も造影集積の高い画像を選択した。入手した DICOM データを jpeg、png に変換し、切り抜き (crop)、分割 (segmentation) などの処理をした。臨床診断は各 US 検査の診断レポートから得た。

データ解析

微小転移 (pN1mi: リンパ節転移の大きさ 0.2mm - 2.0mm) とマクロ転移 (pN1a 以上; リンパ節転移の大きさ < 2.0mm) を、ターゲットラベルデータとして使用した。

解析 (1) 造影 US 画像による深層学習の診断能①

画像の切り抜き

8 年目の乳腺外科医師により 1 枚の造影 US の検査画像から、非造影画像と造影画像を切り抜いた (crop) (図 11)。

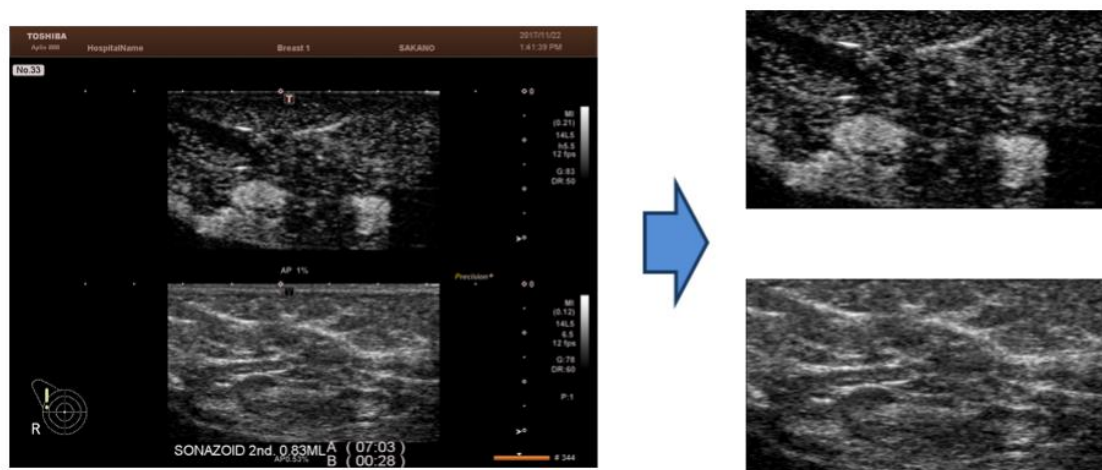


図 11. 1 枚の造影 US 検査画像からの切り抜き (crop)

左がオリジナルの検査画像。右上が造影画像、右下が非造影画像のみ切り抜いた画像

症例

2013 年 1 月～2018 年 12 月までに当科で早期乳癌手術をした症例のうち、術前薬物療法を施行した症例、pStage 0 を除外した 386 例。

陽性症例が 94 例で、陰性症例を残りのうちからランダムに選択して合計 187 例とした。

画像の data augmentation (データ拡張)

通常、深層学習のためのデータ画像数は膨大になるため、臨床で得られるデータだけでは圧倒的に不十分となる。そのため、data augmentation という手法が用いられる。

1. Affine 変換 (Shift scale rotate 法)

Affine 変換とは、Python を用いて画像を平行移動(shift)、拡大(scale)、回転(rotate)等することにより、データ数を増やすことができる。それぞれに Limit 値を設定する。今回は Shift_limit 0.0625, scale_limit 0.1, rotate_limit 10 とした。また、処理施行確率は 0.8 とした。

2. Pairing image 法

2つの異なる症例の画像を重ね合わせて1つの合成画像を作り、正解の数値(ラベル)もその平均をとる手法である(図12)。処理施行確率は0.8とした。

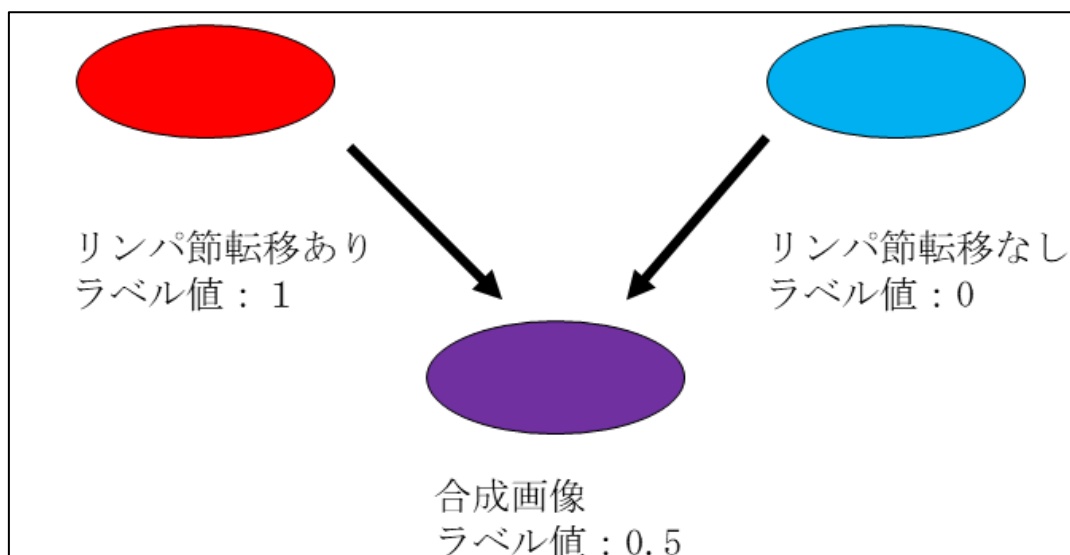


図 12. Pairing image 法 概要図

3. Generative adversarial network (GAN: 敵対的生成ネットワーク)

敵対的生成ネットワーク (GAN)は Generator (生成ネットワーク)と Discriminator (識別ネットワーク)の2つのネットワークから構成されており、互いに競い合わせることで精度を高める Augmentation の1つである。

Auxiliary classifier generative adversarial nets (ACGAN) という、生成画像がラベル情報を保持するように改良された GAN を利用した。

事前学習には The RSNA pulmonary embolism CT dataset (Colak E et al, 2021)を用いた。

深層学習モデル設定

- EfficientNet B0 (Tan M and Le QV, 2019)を利用した。
- Stratified K-fold 法、k=5 でトレーニングコホート、テストコホートに分けて性能を評価した。

検証① Affine 変換と Pairing image 法

- ハイパーパラメータ：最適化関数は Adam (Kingma DP and Ba J, 2014), 学習率 0.00001, バッチサイズ 64, 学習ステップ 2000
- Augmentation 無し、Affine 変換のみ、Pairing Image 法のみ、Affine 変換+Pairing image 法の4種類を比較した。

検証② GAN の追加

- ・ Pairing Image 法をベースに、GAN による augmentation も追加した。
- ・ ハイパーパラメータ：最適化関数は Adam (Kingma DP and Ba J, 2014), 学習率 0.0002、バッチサイズ 100、事前学習ステップ 300000
- ・ 転移学習ステップを 10、100、1000、10000 で比較した。

解析 (2) 造影 US 画像による深層学習の診断能②

◎解析①の反省

新仮説 1：枚数が少ない可能性 (訓練 149 枚、テスト 38 枚)

データ数を増やせば診断能が向上する可能性を考慮した。これにより可能な限りの造影 US 画像施行症例を集め、788 例となった。

新仮説 2：セグメンテーションがあれば診断能が向上する

人間に判別可能ではない解析を人工知能に行わせると、必要枚数を押し上げる因子になりうる。判別するためのわかりやすい特徴があれば、それを見つけ出して一般化してくれるのが人工知能である。一般化しないということは、訓練画像とラベル (予測したいこと) を結びつける特徴に乏しいことを意味する。

→ 医師による手入力で、リンパ節の部分をセグメンテーションした。

症例 (図 13)

対象症例

2013 年～2021 年に当院で早期乳癌の手術を受けた患者

除外症例

(a) 造影 US で腋窩リンパ節を評価しなかった。

※ 術前薬物療法を受けた場合は、術前薬物療法後の造影 US で評価した。

(b) 腋窩手術が行われなかった。

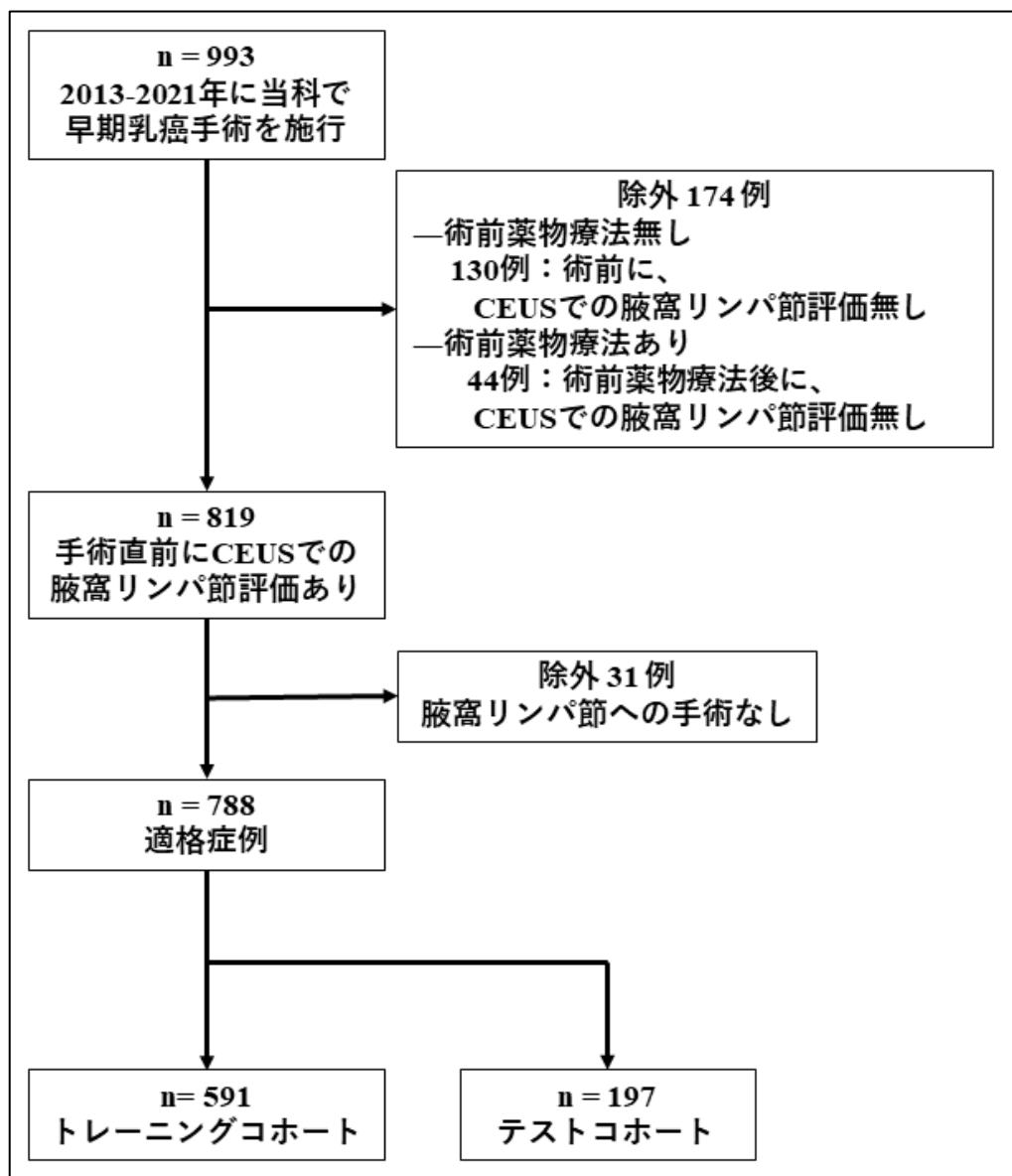


図 13. メイン解析のフローチャート

学習条件

- EfficientNet V2 B0 で比較
- トレーニング:テスト = 3:1 で分割
- 入力画像サイズ：512×512 pixel, 768×768pixel の 2 種類を検証した
- 入力画像の種類 (図 14)
 - (1)pe 非造影 US と造影 US の画像を入力 (赤と緑のみ)
 - (2)pem 非造影 US と造影 US とマスクの画像を入力 (赤と緑と青)
 - (3)crop セグメンテーションでリンパ節だけクロップした画像を入力

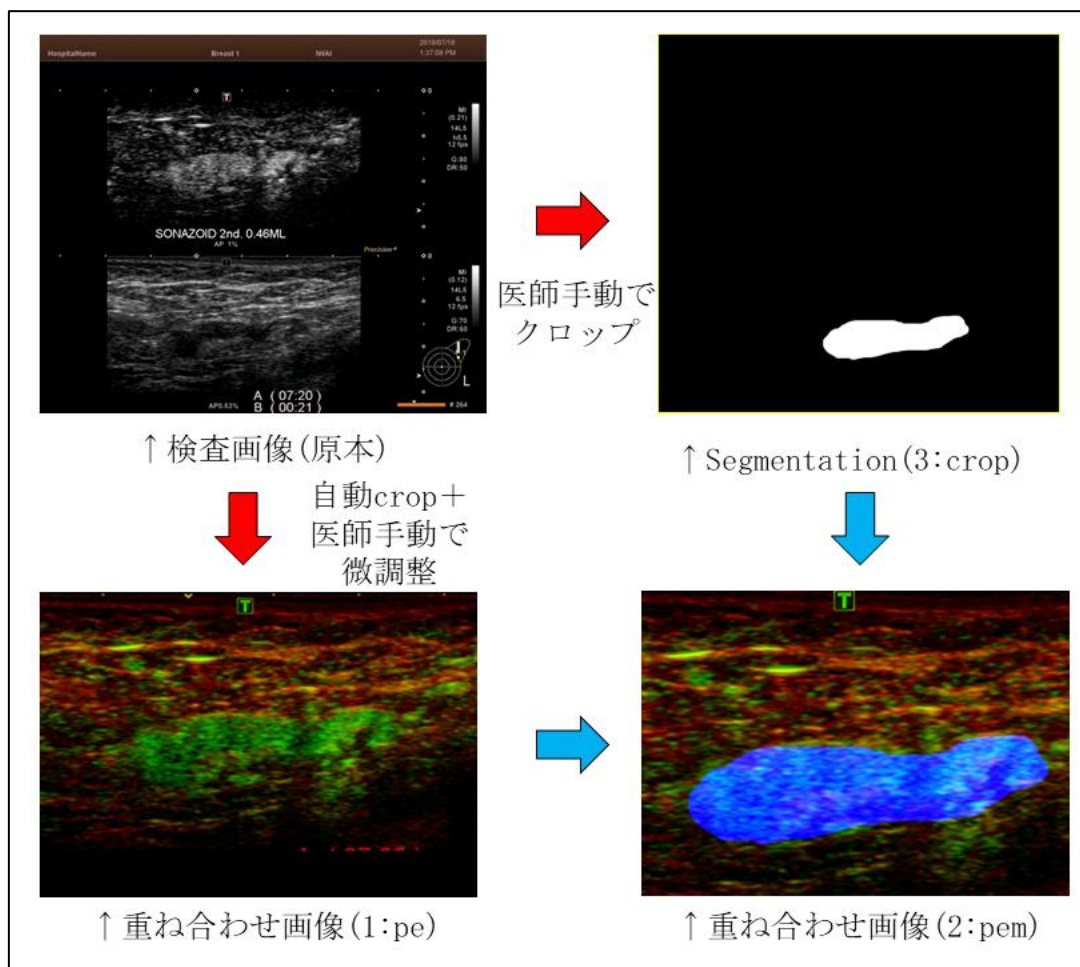


図 14. 解析 2 における、入力画像の処理

赤：非造影 US での高エコー域 緑：造影 US での造影範囲

白、青：医師手入力のリンパ節領域 segmentation

→2 (画像サイズ) x 3 (画像の種類) の 6 条件で検証した

最終解析

症例

解析②と同じである

人工知能解析の概要

①画像による深層学習解析と、②表形式データに①のデータも合わせた機械学習解析の 2 つの解析がある (図 15)

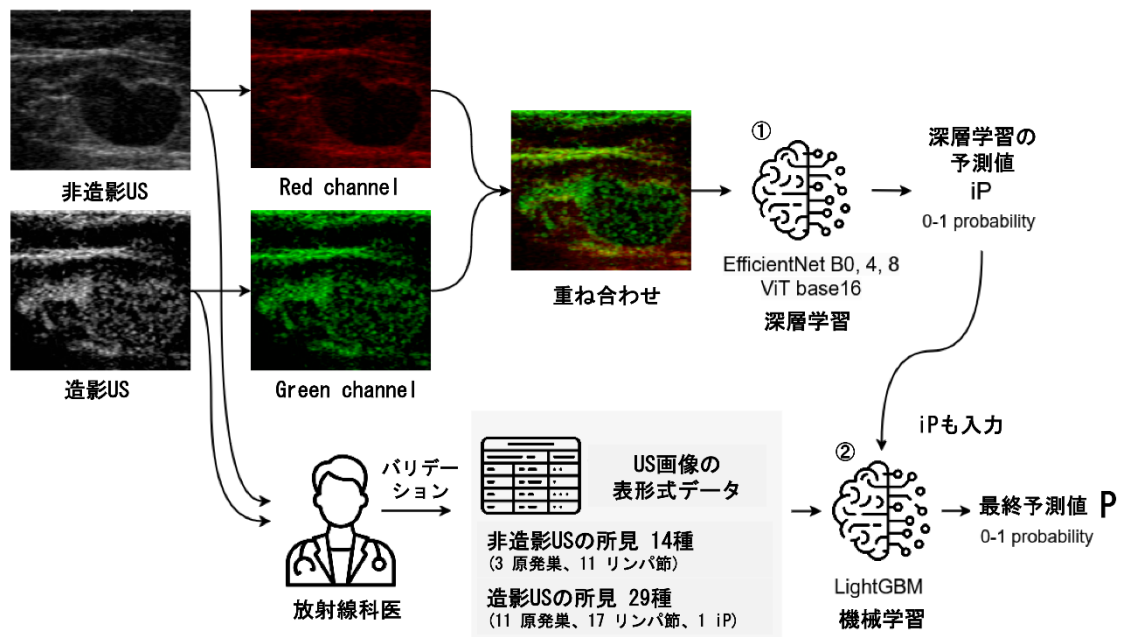


図 15. 本解析での人工知能解析

本研究には 2 つの解析が含まれている。

- ①深層学習モデルに非造影 US と造影 US 画像を重ね合わせた画像 (解析 2 の pe 画像と同等)を入力し、画像による腋窩リンパ節転移の予測共変量 (iP:0-1)を出力する。
- ②機械学習モデルを用いて、画像の所見を医師が読影し表形式データ化した表形式データと深層学習の予測値 (iP)を組み合わせ、最終的な予測値を出力した (iP の追加は、疑似的に医師の診断に深層学習の診断も参考に加えるようなものである)。さらに、特徴量の重要度も評価した。

①深層学習モデル

画像の前処理

非造影 US 画像と造影 US 画像のペアの画像から関心領域が切り取られ、512×512 ピクセル (VisionTransformer モデルでは 384×384 ピクセル)にリサイズされ、平均色と標準色に基づいて正規化された。

両画像は ITU-R BT.709 係数 ($Y = 0.2989R + 0.5870G + 0.1140B$)を用いてグレースケールに変換された。最終的な RGB 合成画像では、グレースケールの非造影 US 画像が R チャンネルに、グレースケールの造影 US 画像が G チャンネルに割り当てられ、B チャンネルはゼロに設定された。これらの合成画像はその後、画像ベースの予測共変量 (iP: 0-1)を導出するために、低い処理コスト (Vu MH et al, 2020)で深層学習解析に使用された。

深層学習解析の詳細設定

以下の条件により、深層学習解析を行った。

使用モデル

①EfficientNetB0、B4、B8 (Tan M and Le QV, 2019)

②ViT Base/16 (Dosovitskiy A et al, 2020)

事前学習

いずれのモデルでも ImageNet (Jia D et al, 2009)、PyTorch 画像モデルライブラリ (Wightman R, 2019)を事前学習に利用した。

画像拡張 (Image augmentation)

過学習やサンプル数の不均衡を緩和するため、ランダムなサイズ変更、クロッピング、水平反転、ノイズの追加、色の配置 (Buslaev A et al, 2020)などを適用した。

ハイパーパラメータ

トレーニングコホートはさらにトレーニングセットとバリデーションセットに 4:1 の比率で分割された。バリデーションセットを使用して最適なハイパーパラメータを決定した後、トレーニングコホート全体を使用して最終モデルをトレーニングした。

各モデルは、シグモイド活性化関数を用い、損失関数としてバイナリー・クロスエントロピーを用いて学習された。学習は Adam optimizer (Kingma DP and Ba J, 2014)を使用し、学習率 $1e-3$ 、バッチサイズ 2、エポック 30 回で行った。

②機械学習モデル

表形式データ

以下のように非造影 US や造影 US の画像データを表形式データに変換し、機械学習に利用した。

乳癌原発巣

腫瘍径・BI-RADS カテゴリー (Carl J DO et al, 2013)

Category 1: 所見なし、Category 2: 良性、Category 3: おそらく良性、

Category 4: 悪性疑い、Category 5: 悪性を強く疑う

造影所見 (Xiao X et al, 2014)

腋窩リンパ節

短径・長径・リンパ節の形態的特徴の t1-9 (Zheng X et al, 2020)、造影パターンに関する特徴の B1-B6 (Omobolaji Alabi R et al, 2022) (表 6)

定性的特徴は 0 か 1 の二変量または 0-7 のカテゴリ変量に変換した。

これらの特徴は、経験年数 9 年と 34 年の 2 人の乳腺外科専門医によって評価され、放射線科医により裏付けされた。9 年の乳腺外科医が全例を評価し、34 年の乳腺外科専門医は無作為に選んだ 80 人の患者を評価観察者間の一致を検証した (カップパー検定)。

表 6. LightGBM に入力した非造影 US / 造影 US による特徴量

非造影 US、乳癌原発巣	
	BI-RADS カテゴリー
	浸潤部の最大径
造影 US、乳癌原発巣	
	浸潤部の最大径
造影所見 1-9	
1	周囲乳腺よりも早期の造影
2	周囲乳腺よりも強い造影: 明らかに強い場合
3	造影の方向 (周囲から/中心から/びまん性)
4	不均一な造影
5	造影領域の境界不明瞭
6	造影の不整な形状
7	リング状の造影効果
8	造影範囲が非造影 US での病変よりも大きい
9	造影欠損
非造影 US、腋窩リンパ節	
	リンパ節の長径、短径
形態所見 t1-t9	
t1	短径/長径比 <2
t2	びまん性の皮質肥厚 >3 mm
t3	局所的な皮質肥厚 >3 mm
t4	求心性の皮質肥厚 >3 mm
t5	リンパ門の消失
t6	円形の低エコー病変
t7	腫瘍によるリンパ節の置換
t8	リンパ節内の微小石灰化
t9	リンパ門以外からの血流の流入
造影 US、腋窩リンパ節	
	リンパ節の長径、短径
	造影 US で評価した t1-t9
造影所見 B1-B6	
B1	求心性の造影
B2	不均一な造影
B3	造影欠損
B4	リング状の造影効果
B5	造影範囲の非造影 US リンパ節からの拡大

LightGBM と特徴選択

これらの特徴量を利用し、LightGBM モデル (Ke G, 2017)により機械学習の解析を行った。すなわち、2 種類の特徴量がある：深層学習による予測 iP (0-1)と表 6 に示す画像の特徴である。出力は腋窩リンパ節転移の予測値 0-1 であった。

iP と造影 US、非造影 US それぞれの特徴を組み合わせることで予測値を算出し、その診断精度を評価した。

トレーニング中、表形式データの特徴の重要度が定量化された。

学習パラメータ設定

リーフ数 (number of leaves) =31、最大深さ (maximum depth) =3、リーフ当たりの最小データ数 (minimum data in leaf) =20、学習率 (learning rate) =0.1、早期終了回数 (early stopping round) =50 とした。

検証は過学習を最小化するために 5-fold cross validation 法を用いてトレーニングコホートで行われた。

重要度の高い所見のみによる診断能

診断を簡略化するため、LightGBM による分析から得られた重要度の高い 3 つの所見を選択し、これらのみを用いて診断能を評価した。テストコホートを用いて、腋窩リンパ節転移予測の正診率、感度、特異度を、所見の陽性所見の数に応じてロジスティック解析により計算した。

診断能評価

診断能の評価指標は、以下の通りに計算している。

感度 = (真陽性) / (真陽性 + 偽陰性)

特異度 = (真陰性) / (真陰性 + 偽陽性)

陽性適中率 = (真陽性) / (真陽性 + 偽陽性)

陰性適中率 = (真陰性) / (真陰性 + 偽陰性)

精度 = (真陽性 + 真陰性) / (全症例)

統計解析

すべての統計は両側で行い、 p 値が 0.05 未満の場合を統計的有意性ありとした。トレーニングセットとテストセットにおける臨床病理学的因子の比較には、t-test か Mann-Whitney U test を使用した。2 つの ROC 曲線の診断能を比較する際には、DeLong test を利用した。US 所見の診断者間の一致の評価には、カッパーテストを利用した。すべての解析は BellCurve for Excel 4.04 または R 4.3.1.を用いて行った。

ソフトウェア

深層学習 : Python 3.12.8 and PyTorch 2.3.0.

機械学習 : LightGBM 4.3.0.

利用可能なコード

以下に実装コードが公開されている。

深層学習 : <https://github.com/endaaman/br-image>

LightGBM : <https://github.com/endaaman/br-clinical>

結果

解析①

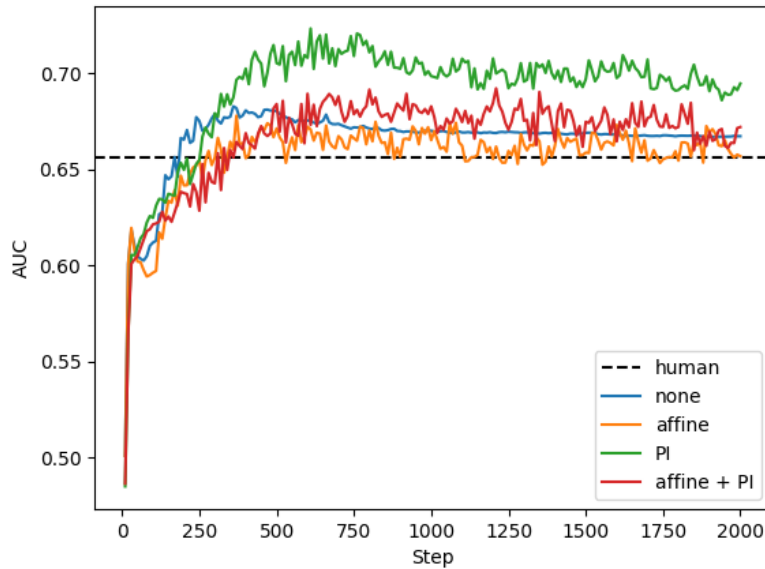


図 16. 解析①の実験 1

Augmentation 無し、Affine 変換のみ、Pairing Image (PI)法のみ、Affine 変換+Pairing image 法の組み合わせの 4 種類を比較した。Pairing Image 法のみが、最も優れた AUC を得られた。

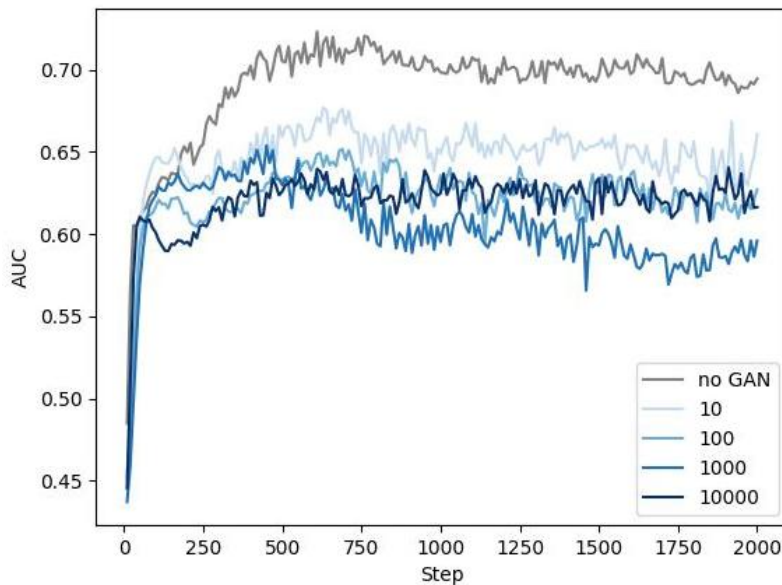


図 17. 解析①の実験 2

Pairing Image 法をベースに、GAN による augmentation も追加した。転移学習ステップを 10,100,1000,10000 で比較した。GAN を追加しない場合に、最も高い AUC を得ることができた。

解析②

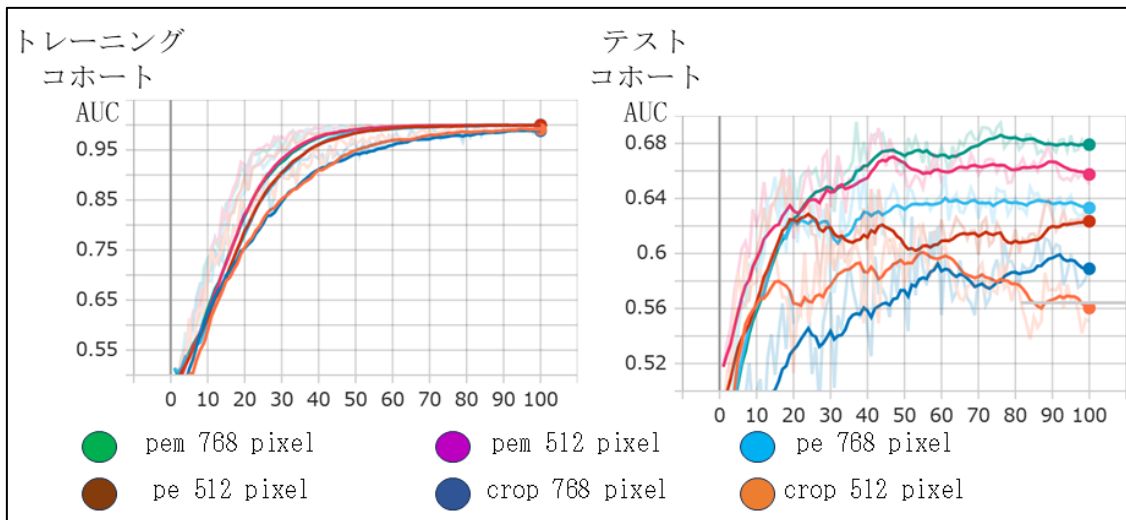


図 18. 解析②における深層学習の腋窩リンパ節転移診断能

(左):トレーニングデータでの ROC

トレーニングデータの範囲内では比較的良好な ROC 曲線であり、特に pe や pem で良い結果となった。

(右):テストデータでの ROC.

テストデータによる有用性を証明できず、本解析の有用性は証明できなかった。

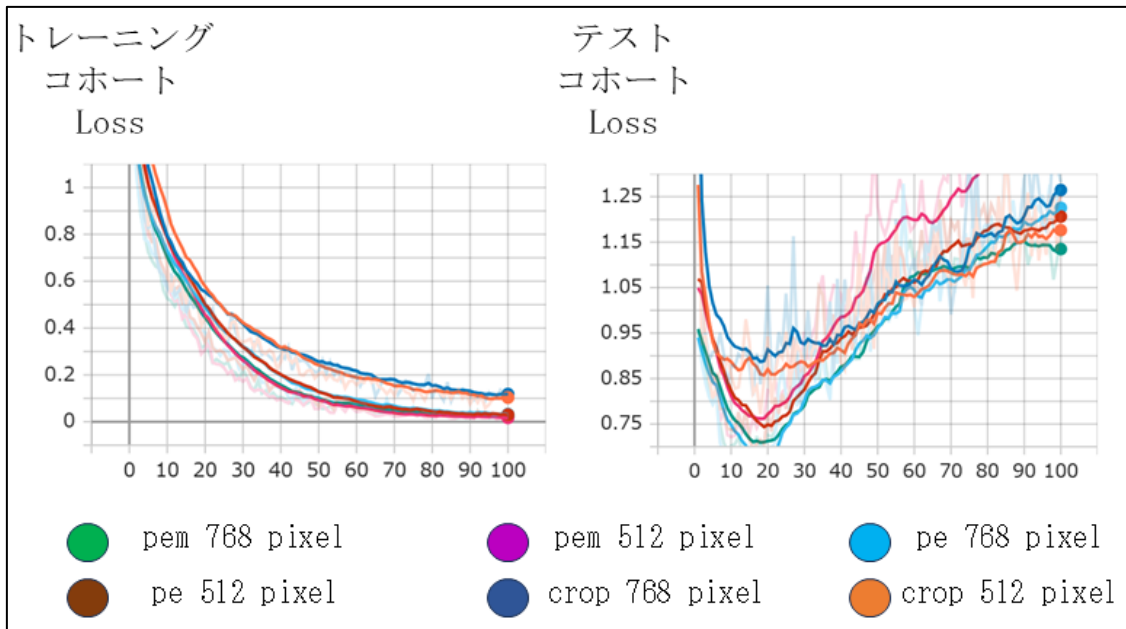


図 19(左):トレーニングデータでの Loss 関数 (右):テストデータでの Loss 関数

Loss 関数の Loss が低下することが訓練の証明となり、パラメータを調整し最適化を行う。テストコホートでは Loss が低下しておらず、最適化されていないと判断され、上記 AUC と矛盾はしなかった。

結果小括

- ・入力サイズは 768 pixel が 568 pixel よりも良好な診断能だった。
- ・入力画像の種類は pem > pe > crop の順で良い結果となった。
- ・訓練コホートで約 100%、テストコホートで約 70%の診断能であり、過学習気味であった。
- ・総じて、報告するには精度不足であった。

本解析

症例

全症例の基本情報の内訳は以下のとおりである。

臨床病期

cStage 0 : 182 例 (23%)

cStage 1 : 331 例 (42%)

cStage 2 : 247 例 (31%)

cStage 3 : 28 例 (4%)

腋窩リンパ節転移

転移無し(pN0) : 598 例 (76%)

転移あり ; 190 例 (24%)

リンパ節転移の内訳

センチネルリンパ節生検または腋窩郭清により摘出された手術標本における、腋窩リンパ節の病理診断

—微小転移のみ (0.2mm 以上、2mm 未満, pN1mi) : 39 例 (5%)、

—マクロ転移を含む 1~3 個の転移 (pN1a) : 111 例 (14%)、

—4~9 個の転移 (pN2) : 30 例 (4%)

—10 個以上の転移 (pN3) : 10 例 (1%)

コホート分け

合計 788 例の患者がトレーニングコホート(591)とテストコホート(197)に 3:1 で分けられた。それぞれのコホートでの臨床病理学的因子を表 7 に示す
臨床病理学的因子は HER2 だけ差がしたが、影響は無いと考えた。

表 7. 症例の臨床病理学的因子

特徴	トレーニングコホート (n=591)	テストコホート (n=197)	p 値
年齢 (歳) ± 標準偏差	59.6 ± 0.53	59.9 ± 0.87	0.74
非造影 US 浸潤径 (mm) ± 標準偏差	13.4 ± 0.50	15.0 ± 0.94	0.17
BI-RADS カテゴリー			0.33
4A	157 (27%)	57 (29%)	
4B	92 (16%)	31 (16%)	
4C	99 (17%)	41 (21%)	
5	154 (26%)	39 (20%)	
術前全身薬物療法			0.11
あり	144 (24%)	54 (27%)	
エストロゲン受容体			0.32
陽性	471 (80%)	166 (84%)	
プロゲステロン受容体			0.55
陽性	416 (70%)	141 (72%)	
HER2			0.005†
陽性	103 (17%)	23 (12%)	
Ki-67 (%) ± 標準偏差	27.6 ± 0.98	27.8 ± 1.67	0.81
乳癌組織型			0.19
非浸潤性乳管癌	65 (11%)	20 (10%)	
浸潤性乳管癌	436 (74%)	140 (71%)	
浸潤性小葉癌	31 (5%)	16 (8%)	
その他	38 (6%)	16 (8%)	

造影 US 画像を用いた腋窩リンパ節転移予測の、深層学習モデルによる性能比較

EfficientNet B0/B4/B8、および ViT Base/16 を使用して、造影 US 画像を入力として腋窩リンパ節転移を予測する際の各モデルの性能をテストコホートで比較した (表 8)。EfficientNet B8 の AUC (0.69) が最も高く、EfficientNet B8 の出力を②の機械学習にも使用した。

表 8：テストコホートでの深層学習モデルの腋窩リンパ節転移診断能

造影 US 画像を入力画像とし、各種深層学習モデルによる腋窩リンパ節転移の診断能を評価した。EfficientNet B8 が最も高い AUC だった。

	AUC	精度	感度	特異度
EfficientNet B0	0.667	0.679	0.553	0.727
EfficientNet B4	0.685	0.637	0.642	0.638
EfficientNet B8	0.687	0.716	0.553	0.776
ViTbase16	0.673	0.768	0.404	0.895

機械学習での腋窩リンパ節転移の予測

全コホートにおける表形式データセットの所見評価のカッパ値は、観察者間一致で 0.96 (0.94-0.97)であった ($p<0.001$ 、カッパ検定)。

テストコホートにおける LightGBM モデルでの診断能を表 9 に示す。精度・感度・特異度は Youden index における診断能である。特に AUC は造影 US を追加した場合に 0.93 (0.88-0.98)と、非造影 US 単独の 0.76 (0.68-0.85)よりも高かった ($p<0.001$, DeLong test) (図 20)。

表 9：テストコホートでの機械学習における診断能

	AUC	精度	感度	特異度
造影 US+iP+非造影 US	0.93 (0.88-0.98)	0.93	0.88	0.95
造影 US	0.93 (0.87-0.98)	0.92	0.88	0.95
iP	0.64 (0.55-0.73)	0.63	0.60	0.64
非造影 US	0.76 (0.66-0.85)	0.72	0.67	0.73

非造影 US と造影 US の診断レポートによる放射線科医の診断

各 US 検査の読影レポートに基づいた全コホートでの診断能を検出した。非造影 US の診断能は精度が 0.82、感度が 0.36、特異度が 0.93 であった。造影 US の診断能は精度が 0.85、感度が 0.45、特異度が 0.95 であった (図 20)。

代表的な画像として、リンパ節転移陽性の診断が一致したものと、機械学習診断が放射線科医の診断を助けるような画像を図 21 に示す。

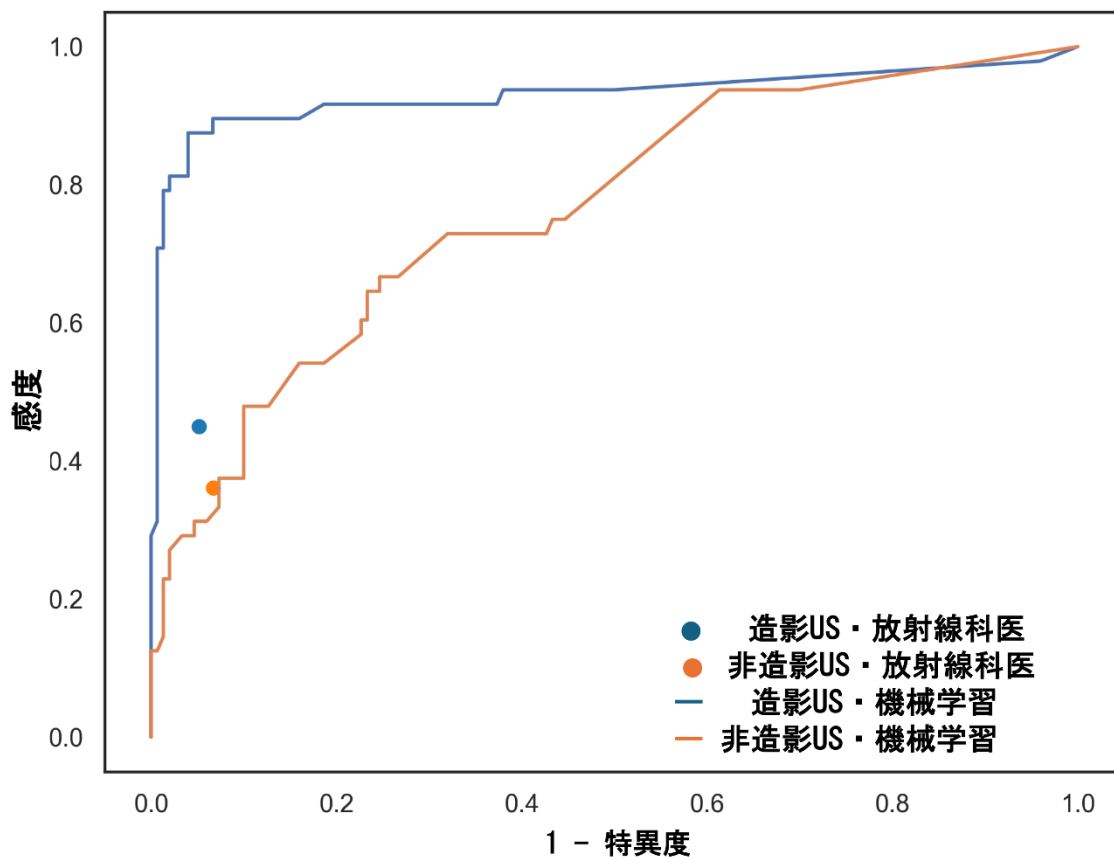


図 20. 腋窩リンパ節転移予測における LightGBM と臨床の診断能

造影 US を追加した場合は、非造影 US のみの 0.76 [0.68-0.85] よりも有意に高い AUC 値 0.93 [0.88-0.98] を示した ($p < 0.001$, DeLong test)。

非造影 US の臨床診断では、感度 0.36、特異度 0.93 であり、造影 US の診断では、感度 0.45、特異度 0.95 であった。

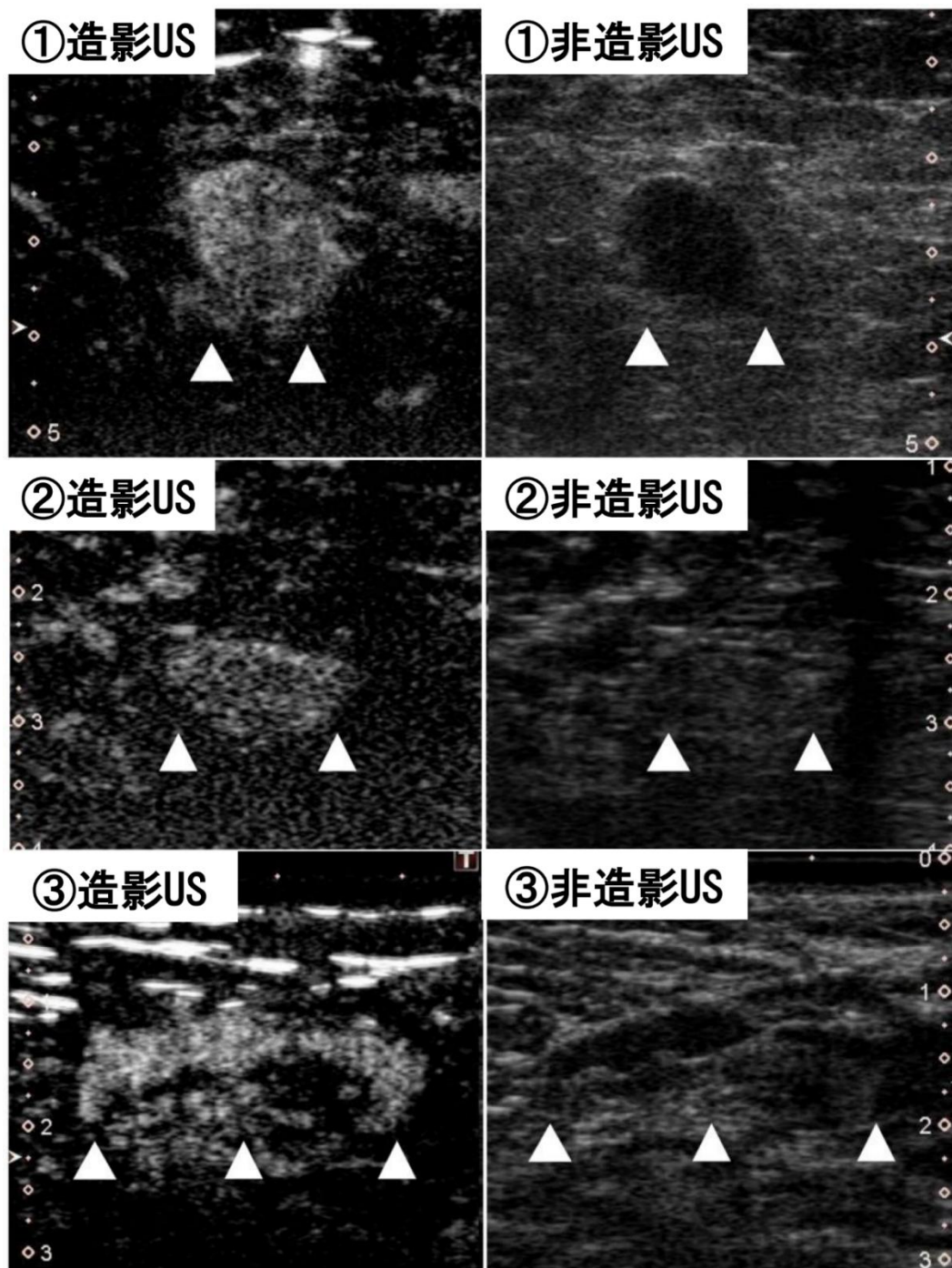


図 21. 各診断の一致性と代表的な腋窩リンパ節画像

- ①リンパ節転移 陽性、放射線科医 陽性、機械学習 陽性 (一致)
- ②リンパ節転移 陽性、放射線科医 陰性、機械学習 陽性 (機械学習により、見逃しを防止)
- ③リンパ節転移 陰性、放射線科医 陽性、機械学習 陰性 (機械学習により、過剰診断を防止)

特徴量 (所見)の重要度

LightGBM で抽出された特徴の重要度を表 10 に示す。人工知能による予測 iP は 5.14×10^1 と 6 番目に重要度が高かった。

表 10. LightGBM が抽出した特徴量の中での重要度

	特徴	重要度
1. B2	不均一な造影 (造影 US)	3.24×10^2
2. t2	びまん性の皮質肥厚>3mm (造影 US)	3.16×10^2
3. t4	偏心性の皮質肥厚 >3mm (造影 US)	2.58×10^2
4	原発巣の浸潤径 (造影 US)	1.54×10^2
5. B5	造影効果が非造影 US でのリンパ節よりも大きい	1.46×10^2
6	iP (造影 US 画像による深層学習予測)	5.14×10^1

重要度の高い所見のみによる診断能

表 10 の結果を踏まえ、B2、t2、t4 を重要な所見として採用した。テストコホートをを用いて診断能を評価し、表 11 にその結果を示す。特に 1 つまたは 2 つの特徴が陽性であった場合の診断能は、臨床的に有用であると考えられた。

表 11. 重要度の高い所見のみによる診断能

LightGBM で抽出した重要度の高い 3 つの特徴量を用い、その陽性所見の数に応じた診断能を、全コホートをを用いて統計解析した。

陽性所見の数	精度	感度	特異度
3	0.80	0.17	1.00
2 以上	0.86	0.43	0.98
1 以上	0.90	0.68	0.95

考察

1.本研究から得られた新知見

本研究の機械学習モデルは、腋窩リンパ節転移を診断するための造影 US における重要な特徴を抽出した。また、この所見は臨床的にも妥当性を有した。

2.先行研究との比較

重要度の高い所見

LightGBM によって抽出された重要な特徴のうち、不均一な強調 (B2)は、過去の報告 (Niu Z et al, 2022; Niu Z et al, 2023)と一致した。逆に欠損 (B3)も転移を評価する因子として報告されているが (Niu Z et al, 2023)、本研究では重要度は低かった。また、深層学習解析に原発巣所見を用いた先行報告 (Zheng X et al, 2020; Zhou LQ et al, 2020)があるため、原発巣所見も含めた。原発巣の浸潤径は4番目に重要な所見であった。しかし、これは進行度に直接関係するので当然である。他の原発巣所見は重要度の高いものは少なく、総じて腋窩リンパ節の所見が腋窩リンパ節転移の診断に関与していた。

造影 US は診断指標が統一されておらず普及率の低いものの、このように人工知能により重要な特徴を抽出できれば、造影 US の普及につながる可能性がある。

人工知能解析の精度上昇に向けて。入力データについて。

本研究では造影 US 画像を入力画像とする深層学習モデルを用いて腋窩リンパ節転移の予測を評価した。造影 US 画像のみの深層学習モデルによる診断 iP は、LightGBM の診断において重要度は6番目であり高い重要性を示せなかった。この理由として、US 検査の主観性と、入力画像と実臨床での評価項目の不一致が理由として考えられた。

乳癌においても US 検査には術者依存性があり、標準化は他の画像モダリティよりも困難である (Togawa R et al, 2024; Chung HL et al, 2022)。これらの問題に対処するため、腋窩リンパ節転移を非造影 US で診断するための明確で標準化されたプロトコールと基準を用いた臨床試験が最近開始された (Chang JM et al, 2021)。このように、統一された画像診断による US 画像が得られれば、人工知能の診断能も向上する可能性がある。

別の要因として、本研究での各症例における入力画像は、非造影 US 画像と造影 US 画像を重ね合わせた1枚のみである。一方、通常の臨床的な読影では造影 US の評価所見には非造影 US との比較や、造影剤の経時的変化なども診断基準に含まれる。しかし、本研究では経時変化などの情報は含まれていなかった。これが、診断能が不十分であった理由の1つと考えられる。実際に、ピーク強度 (Du LW et al, 2021; Zhu AQ et al, 2020)、ピークまでの時間 (Du LW et al, 2021)、造影持続時間 (Du LW et al, 2021)などの造影剤の経時的な所見は腋窩リンパ節診断に寄与する。当院

での造影 US の検査において、2 回の造影剤投与で腋窩リンパ節をメインの造影剤の経時的変化を評価対象とすることは、乳房に関心病変が無い場合に限られ、大半の症例が late phase での評価のみであった。そのため本研究では造影の時間経過の解析は行わなかった。近年では、甲状腺 US の静止画像と動画を深層学習モデルへの入力データとして用いた報告があり、静止画像よりも動画の方が診断能において優れていた (Zhang C et al, 2022)。このように入力データとして動画や複数の画像を用いることで、画像選択によるバイアスをなくしながら、経時的変化の所見も加わって診断能が向上する可能性がある。

人工知能解析の精度上昇に向けて—人工知能モデルについて—

本研究では旧来の深層学習 モデルである Convolutional neural network モデルの他に Vision transformer モデルの診断能を検証した。これは ChatGPT®などにも利用される自然言語用の Transformer モデルを画像解析用に変換したモデルであり (Dosovitskiy A et al, 2020)、Convolutional neural network の診断能を超える診断能を持つという報告も見られる (Ayana G et al, 2023)。しかし、本研究では同等の結果しか得られなかった。Vision transformer モデルには Convolutional neural network 以上にデータ数が必要という見解もあり、本症例でもデータ数が不足していた可能性がある。データ数を増やし最新の深層学習モデルに対応できれば、診断能が向上する可能性もある。

臨床活用の可能性

本研究で LightGBM を用いて抽出された重要な特徴による腋窩リンパ節転移の診断は、臨床的に有用である可能性がある。腋窩リンパ節転移の評価は早期乳癌の治療方針を決定する上で重要であるものの、診断能だけでなく検査の負担も考慮する必要がある。造影 US を利用してセンチネルリンパ節生検、マーカー植え込み、直接の針生検などを行い、腋窩リンパ節転移の診断を行う報告もあり (Du LW et al, 2021; Nielsen Moody A et al, 2020)、これらは高い診断能が報告されているものの、侵襲的な手技であり検査には時間も要するため、全例に実施することは困難である。

一方、深層学習を用いた腋窩リンパ節転移の診断には、造影 CT (Liu Z et al, 2021) や造影 MRI (Gao J et al, 2023)を用いた報告もあり、専門医の診断能を凌駕している。診断性能は高くても、これらの検査は US 検査に比べ高価で利便性が低く、放射線被曝や侵襲性などのリスクも伴う (Marino MA et al, 2020)。腋窩リンパ節転移の画像診断に造影 US を用いれば、US の利便性を生かしつつ、さらに高い診断能を得ることが可能であり、治療戦略に貢献する可能性がある。

3.本研究の問題点

本研究の特徴は、検索し得る限り最多の造影 US の症例数で行え、造影 US 画像による深層学習の先駆けとなったことが挙げられる。機械学習による診断と所見の重要性の評価は、高い診断能と実臨床での診断簡略化の可能性を示し、人工知能の解析を臨床に取り入れる一例となる可能性がある。一方で以下の問題点もあり、今後の課題となる。

①単施設後ろ向きである

本研究では外部の検証コホートが含まれていなかった。そのため、選択バイアスや想起バイアスがあった可能性がある。モデルの性能を検証するために、複数の施設で前向きにさらなる研究を行うべきである。

②過学習の可能性

機械学習ではトレーニング中に 5 重クロスバリデーションが実施され、深層学習ではデータ拡張が行われたものの、モデルが過学習を起こした可能性がある。今後はより大きなデータセットを分析する必要がある。

③特徴量の主観性

機械学習における特徴量の評価が医師の判断によって決定されており、観察者バイアスとなる可能性がある。とりわけ人工知能研究において重要なエンドツーエンド (end-to-end) の研究ではなかったことを示しており、モデル全体の最適化が制限された可能性もある。

全体の結論

①本研究から得られた知見

第1章

- (1)非造影 US と造影 US との乳癌原発巣の大きさの相違 (Discrepancy)は、腋窩リンパ節転移の独立した関連因子であった。
- (2)Discrepancy は無再発生存期間などの予後とも関連していた。

第2章

- (1)深層学習モデルを用いた造影 US 画像による早期乳癌の腋窩リンパ節転移予測は、機械学習診断の補助になり得るものの、改善の余地がある
- (2)機械学習を用いた解析は、造影 US の追加により高い診断能を得ることができた。さらに”不均一な造影効果”や”びまん性の皮質肥厚”, “求心性の皮質肥厚”などの重要な特徴量を抽出することができた。

②新知見の意義

第1章

乳癌の診療に重要である腋窩リンパ節転移を予測し得る、簡便な指標が1つ得られ、より最適な乳癌診療につながる可能性がある。

今回の知見は、単に診断指標が増えただけにとどまらない。腋窩リンパ節の診断に、腋窩リンパ節ではなく乳癌原発巣の評価が利用し得たこと、さらに単一の画像検査だけでなく、造影 US と非造影 US の所見の組み合わせが診断に寄与したことは、腋窩リンパ節転移の診断に対する新たなアプローチへと続く可能性を秘めている。

さらに本研究全体として、人工知能を導入する前に通常診療の範囲内でも新たな知見は得られたことは、人工知能に頼らずとも日常診療はまだ改善の余地を残していることを示唆している。

第2章

- (1)造影 US 画像を用いて深層学習による解析が可能である。また、深層学習の診断を機械学習の診断に組み合わせることが可能である。
- (2)機械学習により抽出した特徴量(所見)は、これまでの既報と比較することでその所見の有用性の裏付けまたは反証となり得る。

③本研究で得られた新知見から、今後どのような研究が展開されうるか

第1章

Discrepancy を実臨床に応用するために、造影 US 検査を臨床に取り入れる多施

設共同前向き臨床試験により、その実用性を評価する研究が求められる。この際に、第2章で得られた腋窩リンパ節の重要な所見 “不均一な造影効果”、“びまん性の皮質肥厚”、“求心性の皮質肥厚”などを組み合わせることも考慮される。

第2章

造影 US の動画を深層学習モデルに学習させ、腋窩リンパ節転移の診断能の評価を行う研究が考えられる。また、学習させた深層学習モデルで腋窩リンパ節転移の有無を予測させ、それを読影医に参考にさせた場合の読影結果への影響の評価を行う研究も考えられる。同様の研究は機械学習モデルでも考えられる。

また機械学習により抽出した特徴量や第1章の **Discrepancy** を追加し、乳癌の腋窩リンパ節転移の有無を評価し、それにより治療方針を決定しその予後を比較する、前向き臨床研究も考えられる。

④今後の課題

第1章

Discrepancy の評価自体の課題としては、本研究で対象外とした腫瘍径が 20mm 未満の症例や 50mm を超える症例にも、適切な閾値とともに適応できるように検証することが挙げられる。さらに超音波検査の主観性に対応するために、より明確な診断基準を模索することが課題として挙げられる。

Discrepancy を実臨床に応用するための課題には、特に造影 US 施行の時間的なコストが挙げられる。臨床試験へつなげるには、腋窩リンパ節転移の画像診断が重要な症例に限定することが必須と考えられる。

第2章

第1章でも述べた超音波検査の主観性は、機械学習・深層学習共に課題となる。また、深層学習のためには、本研究以上に均一に撮像・処理された画像を、本研究以上に多くの数で求められる。さらに造影 US 動画の活用も求められる。

謝辞

本研究に関して以下の方々に心より感謝申し上げます。

大学院医学院での研究の機会を与えてくださった北海道大学大学院医学院外科学講座乳腺外科学教室の山下啓子前教授並びに、継続でご指導頂いた高橋將人現教授、人工知能の利用についてご提案とご指導をしてくださった竹下卓志先生に深い感謝の意を表します。

人工知能解析について、初回解析は北海道大学大学院情報科学研究科情報理工学専攻情報システム設計学研究室の学士 成木太音様にご施行頂きました。深い感謝の意を表します。

本解析には北海道大学大学院医学研究院 病理学講座腫瘍病理学教室の田中伸哉教授のご指導の下、同教室の大学院生である遠田建先生にご施行頂きました。深い感謝の意を表します。

また、同教室の大学院生であり北海道大学大学院医学研究院 機能再生医学分野 整形外科学教室の医員である清水寛和先生には本研究全体を通じてご指導・ご協力頂きました。深く感謝致します。

画像診断解析や画像データの入手については、北海道大学病院放射線診断科の加藤扶美講師にご協力頂きました。深い感謝の意を表します。

最後に、臨床データを使用させていただいた患者の皆様、当該分野内外の先生方、その他の多くの方々の御指導と御協力に支えられて本論文は完成しました。研究に携わったすべての方々へ心からの感謝の意を表します。

利益相反

本研究に関して、著者ならびに関係者に開示すべき利益相反状態はありません

引用文献

- Ayana G, Dese K, Dereje Y, Kebede Y, Barki H, Amdissa D, Husen N, Mulugeta F, Habtamu B and Choe SW (2023). Vision-transformer-based transfer learning for mammogram classification. *Diagnostics (Basel)*, 13.
- Bhargava A, Popel AS and Pathak AP (2023). Vascular phenotyping of the invasive front in breast cancer using a 3d angiogenesis atlas. *Microvasc Res*, 149, 104555.
- Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M and Kalinin AA (2020). Albumentations: Fast and flexible image augmentations. *Information*, 11, 125.
- Cantisani V, Di Leo N, David E and Clevert DA (2021). Role of ceus in vascular pathology. *Ultraschall Med*, 42, 348-366.
- Carl J DO, EA. S, EB. M and EA. M (2013). *Acr bi-rads® atlas, breast imaging reporting and data system. 5th ed.*, Reston, VA, American College of Radiology.
- Chang JM, Leung JWT, Moy L, Ha SM and Moon WK (2020). Axillary nodal evaluation in breast cancer: State of the art. *Radiology*, 295, 500-515.
- Chang JM, Shin HJ, Choi JS, Shin SU, Choi BH, Kim MJ, Yoon JH, Chung J, Kim TH, Han BK, et al (2021). Imaging protocol and criteria for evaluation of axillary lymph nodes in the nautilus trial. *J Breast Cancer*, 24, 554-560.
- Christofori G (2006). New signals from the invasive front. *Nature*, 441, 444-50.
- Chung HL, Tso HH, Middleton LP, Sun J and Leung JWT (2022). Axillary nodal metastases in invasive lobular carcinoma versus invasive ductal carcinoma: Comparison of node detection and morphology by ultrasound. *AJR Am J Roentgenol*, 218, 33-41.
- Colak E, Kitamura FC, Hobbs SB, Wu CC, Lungren MP, Prevedello LM, Kalpathy-Cramer J, Ball RL, Shih G, Stein A, et al (2021). The rsna pulmonary embolism ct dataset. *Radiol Artif Intell*, 3, e200254.
- Cox K, Sever A, Jones S, Weeks J, Mills P, Devalia H, Fish D and Jones P (2013). Validation of a technique using microbubbles and contrast enhanced ultrasound (ceus) to biopsy sentinel lymph nodes (sln) in pre-operative breast cancer patients with a normal grey-scale axillary ultrasound. *Eur J Surg Oncol*, 39, 760-5.
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. Available: <https://ui.adsabs.harvard.edu/abs/2020arXiv201011929D> <https://arxiv.org/pdf/2010.11929.pdf> [Accessed October 01, 2020].
- Du LW, Liu HL, Gong HY, Ling LJ, Wang S, Li CY and Zong M (2021). Adding contrast-enhanced ultrasound markers to conventional axillary ultrasound improves specificity for predicting axillary lymph node metastasis in patients with breast cancer. *Br J Radiol*, 94, 20200874.

Fares J, Fares MY, Khachfe HH, Salhab HA and Fares Y (2020). Molecular principles of metastasis: A hallmark of cancer revisited. *Signal Transduct Target Ther*; 5, 28.

Fu Y, Zhong J, Tan Y, Zheng T, Liu M and Wang G (2024). Contrast-enhanced ultrasound for differentiating benign from malignant focal solid renal lesions in pediatric patients. *Sci Rep*, 14, 11409.

Fujisawa T, Hirakata T, Yanagita Y, Iijima M, Horikoshi H, Takeuchi K and Saitoh Y (2013). The detection of pcr after pst by contrast-enhanced ultrasonography for breast cancer. *Breast Cancer*; 20, 75-82.

Gao J, Zhong X, Li W, Li Q, Shao H, Wang Z, Dai Y, Ma H, Shi Y, Zhang H, et al (2023). Attention-based deep learning for the preoperative differentiation of axillary lymph node metastasis in breast cancer on dce-mri. *J Magn Reson Imaging*, 57, 1842-1853.

Garcia-Ortega MJ, Benito MA, Vahamonde EF, Torres PR, Velasco AB and Paredes MM (2011). Pretreatment axillary ultrasonography and core biopsy in patients with suspected breast cancer: Diagnostic accuracy and impact on management. *Eur J Radiol*, 79, 64-72.

Gentilini OD, Botteri E, Sangalli C, Galimberti V, Porpiglia M, Agresti R, Luini A, Viale G, Cassano E, Peradze N, et al (2023). Sentinel lymph node biopsy vs no axillary surgery in patients with small breast cancer and negative results on ultrasonography of axillary lymph nodes: The sound randomized clinical trial. *JAMA Oncol*.

Hosny A, Parmar C, Quackenbush J, Schwartz LH and Aerts H (2018). Artificial intelligence in radiology. *Nat Rev Cancer*; 18, 500-510.

Ito T and Komoike Y (2024). Understanding the basics and clinical applications of contrast-enhanced ultrasound for breast lesions. *J Med Ultrason* (2001), 51, 563-566.

Jia D, Wei D, Richard S, Li-Jia L, Kai L and Li F-F (2009). Imagenet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition*, 248-255.

Jiang YX, Liu H, Liu JB, Zhu QL, Sun Q and Chang XY (2007). Breast tumor size assessment: Comparison of conventional ultrasound and contrast-enhanced ultrasound. *Ultrasound Med Biol*, 33, 1873-81.

Jin L, Wang R, Zhuang L, Jin Y, Sun X, Jia C, Lin L, Shi Q, Zhang W and Du L (2022). Evaluation of whole axillary status with lymphatic contrast-enhanced ultrasound in patients with breast cancer. *Eur Radiol*, 32, 630-638.

Johnston SRD, Harbeck N, Hegg R, Toi M, Martin M, Shao ZM, Zhang QY, Martinez Rodriguez JL, Campone M, Hamilton E, et al (2020). Abemaciclib combined with endocrine therapy for the adjuvant treatment of hr+, her2-, node-positive, high-risk, early breast cancer (monarche). *J Clin Oncol*, 38, 3987-3998.

Kalinsky K, Barlow WE, Gralow JR, Meric-Bernstam F, Albain KS, Hayes DF, Lin NU, Perez EA, Goldstein LJ, Chia SKL, et al (2021). 21-gene assay to inform chemotherapy benefit in node-positive breast cancer. *N Engl J Med*, 385, 2336-2347.

Kang YJ, Han W, Park S, You JY, Yi HW, Park S, Nam S, Kim JH, Yun KW, Kim HJ, et al (2017).

Outcome following sentinel lymph node biopsy-guided decisions in breast cancer patients with conversion from positive to negative axillary lymph nodes after neoadjuvant chemotherapy. *Breast Cancer Res Treat*, 166, 473-480.

Ke G (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Adv Neural Inform Process Syst*.

Kingma DP and Ba J. (2014). Adam: A method for stochastic optimization. Available: <https://ui.adsabs.harvard.edu/abs/2014arXiv1412.6980K>
<https://arxiv.org/pdf/1412.6980.pdf> [Accessed December 01, 2014].

Krizhevsky A, Sutskever I and Hinton GE (2012). Imagenet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume I*. Lake Tahoe, Nevada: Curran Associates Inc.

Langer I, Guller U, Berclaz G, Koechli OR, Schaer G, Fehr MK, Hess T, Oertli D, Bronz L, Schnarwyler B, et al (2007). Morbidity of sentinel lymph node biopsy (sln) alone versus sln and completion axillary lymph node dissection after breast cancer surgery: A prospective swiss multicenter study on 659 patients. *Ann Surg*, 245, 452-61.

Liu YB, Xia M, Li YJ, Li S, Li H and Li YL (2021). Contrast-enhanced ultrasound in locating axillary sentinel lymph nodes in patients with breast cancer: A prospective study. *Ultrasound Med Biol*, 47, 1475-1483.

Liu Z, Ni S, Yang C, Sun W, Huang D, Su H, Shu J and Qin N (2021). Axillary lymph node metastasis prediction by contrast-enhanced computed tomography images for breast cancer patients based on deep learning. *Comput Biol Med*, 136, 104715.

Marino MA, Avendano D, Zapata P, Riedl CC and Pinker K (2020). Lymph node imaging in patients with primary breast cancer: Concurrent diagnostic tools. *Oncologist*, 25, e231-e242.

Mills P, Sever A, Weeks J, Fish D, Jones S and Jones P (2010). Axillary ultrasound assessment in primary breast cancer: An audit of 653 cases. *Breast J*, 16, 460-3.

Nielsen Moody A, Cox K, Haigh I, Chen Y and Sharma N (2020). Does contrast enhanced ultrasound (ceus) of normal/benign axillary lymph nodes in patients with breast cancer identify significant axillary nodal burden? *Eur J Radiol*, 132, 109311.

Niu Z, Gao Y, Xiao M, Mao F, Zhou Y, Zhu Q and Jiang Y (2023). Contrast-enhanced lymphatic us can improve the preoperative diagnostic performance for sentinel lymph nodes in early breast cancer. *Eur Radiol*, 33, 1593-1602.

Niu Z, Xiao M, Ma L, Qin J, Li W, Zhang J, Zhu Q and Jiang Y (2022). The value of contrast-enhanced ultrasound enhancement patterns for the diagnosis of sentinel lymph node status in breast cancer: Systematic review and meta-analysis. *Quant Imaging Med Surg*, 12, 936-948.

Novikov NM, Zolotaryova SY, Gautreau AM and Denisov EV (2021). Mutational drivers of cancer cell migration and invasion. *Br J Cancer*, 124, 102-114.

Omobilaji Alabi R, Almangush A, Elmusrati M, Leivo I and Makitie AA (2022). An interpretable machine learning prognostic system for risk stratification in oropharyngeal cancer. *Int J Med*

Inform, 168, 104896.

Ouyang Q, Chen L, Zhao H, Xu R and Lin Q (2010). Detecting metastasis of lymph nodes and predicting aggressiveness in patients with breast carcinomas. *J Ultrasound Med*, 29, 343-52.

Pan J, Guo T, Kong H, Bu W, Shao M and Geng Z (2025). Prediction of mortality risk in patients with severe community-acquired pneumonia in the intensive care unit using machine learning. *Sci Rep*, 15, 1566.

Rafailidis V, Huang DY, Yusuf GT and Sidhu PS (2020). General principles and overview of vascular contrast-enhanced ultrasonography. *Ultrasonography*, 39, 22-42.

Reimer T, Stachs A, Veselinovic K, Kuhn T, Heil J, Polata S, Marme F, Muller T, Hildebrandt G, Krug D, et al (2025). Axillary surgery in breast cancer - primary results of the insema trial. *N Engl J Med*, 392, 1051-1064.

Ribatti D, Tamma R and Annese T (2020). Epithelial-mesenchymal transition in cancer: A historical overview. *Transl Oncol*, 13, 100773.

Rufo DD, Debelee TG, Ibenthal A and Negera WG (2021). Diagnosis of diabetes mellitus using gradient boosting machine (lightgbm). *Diagnostics (Basel)*, 11.

Sanegre S, Eritja N, de Andrea C, Diaz-Martin J, Diaz-Lagares A, Jacome MA, Salguero-Aranda C, Garcia Ros D, Davidson B, Lopez R, et al (2021). Characterizing the invasive tumor front of aggressive uterine adenocarcinoma and leiomyosarcoma. *Front Cell Dev Biol*, 9, 670185.

Sato K, Tamaki K, Tsuda H, Kosuda S, Kusano S, Hiraide H and Mochizuki H (2004). Utility of axillary ultrasound examination to select breast cancer patients suited for optimal sentinel node biopsy. *Am J Surg*, 187, 679-83.

Sidhu PS, Cantisani V, Dietrich CF, Gilja OH, Saftoiu A, Bartels E, Bertolotto M, Calliada F, Clevert DA, Cosgrove D, et al (2018). The efsumb guidelines and recommendations for the clinical practice of contrast-enhanced ultrasound (ceus) in non-hepatic applications: Update 2017 (long version). *Ultraschall Med*, 39, e2-e44.

Siegel RL, Giaquinto AN and Jemal A (2024). Cancer statistics, 2024. *CA Cancer J Clin*, 74, 12-49.

Soerjomataram I, Louwman MW, Ribot JG, Roukema JA and Coebergh JW (2008). An overview of prognostic factors for long-term survivors of breast cancer. *Breast Cancer Res Treat*, 107, 309-30.

Tan M and Le QV. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. Available: <https://ui.adsabs.harvard.edu/abs/2019arXiv190511946T>
<https://arxiv.org/pdf/1905.11946.pdf> [Accessed May 01, 2019].

Togawa R, Pfob A, Busch C, Fastner S, Gomez C, Goncalo M, Hennigs A, Killinger K, Nees J, Riedel F, et al (2024). Intra- and interobserver reliability of shear wave elastography in breast cancer diagnosis. *J Ultrasound Med*, 43, 109-114.

Tutt ANJ, Garber JE, Kaufman B, Viale G, Fumagalli D, Rastogi P, Gelber RD, de Azambuja E, Fielding A, Balmana J, et al (2021). Adjuvant olaparib for patients with brca1- or brca2-mutated

breast cancer. *N Engl J Med*, 384, 2394-2405.

Vamvakas A, Tsivaka D, Logothetis A, Vassiou K and Tsougos I (2022). Breast cancer classification on multiparametric mri - increased performance of boosting ensemble methods. *Technol Cancer Res Treat*, 21, 15330338221087828.

von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, Suter T, Arahmani A, Rouchet N, Clark E, et al (2017). Adjuvant pertuzumab and trastuzumab in early her2-positive breast cancer. *N Engl J Med*, 377, 122-131.

Vu MH, Grimbergen G, Nyholm T and Lofstedt T (2020). Evaluation of multislice inputs to convolutional neural networks for medical image segmentation. *Med Phys*, 47, 6216-6231.

Wang Z, Zhou Q, Liu J, Tang S, Liang X, Zhou Z, He Y, Peng H and Xiao Y (2014). Tumor size of breast invasive ductal cancer measured with contrast-enhanced ultrasound predicts regional lymph node metastasis and n stage. *Int J Clin Exp Pathol*, 7, 6985-91.

Wei Q, Qian Y, Yu J and Wong CC (2020). Metabolic rewiring in the promotion of cancer metastasis: Mechanisms and therapeutic implications. *Oncogene*, 39, 6139-6156.

Wightman R. (2019). *Pytorch image models* [Online]. GitHub. Available: <https://github.com/huggingface/pytorch-image-models> [Accessed August 19, 2023 2023].

Xiao X, Ou B, Yang H, Wu H and Luo B (2014). Breast contrast-enhanced ultrasound: Is a scoring system feasible? A preliminary study in china. *PLoS One*, 9, e105517.

Zhang C, Liu D, Huang L, Zhao Y, Chen L and Guo Y (2022). Classification of thyroid nodules by using deep learning radiomics based on ultrasound dynamic video. *J Ultrasound Med*, 41, 2993-3002.

Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, Mao R, Li F, Xiao Y, Wang Y, et al (2020). Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun*, 11, 1236.

Zhou LQ, Wu XL, Huang SY, Wu GG, Ye HR, Wei Q, Bao LY, Deng YB, Li XR, Cui XW, et al (2020). Lymph node metastasis prediction from primary breast cancer us images using deep learning. *Radiology*, 294, 19-28.

Zhu AQ, Li XL, An LW, Guo LH, Fu HJ, Sun LP and Xu HX (2020). Predicting axillary lymph node metastasis in patients with breast invasive ductal carcinoma with negative axillary ultrasound results using conventional ultrasound and contrast-enhanced ultrasound. *J Ultrasound Med*, 39, 2059-2070.