184　機械学習による新規医療データ統合解析システムの構築	河原　大輔

【目的】本研究の目的は、子宮体がんの脈管侵襲予測モデルの作成および検証を行い、対応する組織の組織染色画像より得られる血管内皮細胞およびがん上皮細胞の形態から Radiomics 特徴量として表現される予測因子と組織学情報との関連を評価することで、脈管侵襲予測に寄与するがん微小環境構造内の特徴を解明することである。

【方法】子宮体がんの脈管侵襲予測モデル構築は、2014 年から 2020 年に広島大学病院にて子宮体がん摘出手術を受けた 153 症例を対象として行った。さらに、2021 年から 2022 年に広島大学病院にて子宮体がん摘出手術を受けた 32 症例を対象として予測モデルの検証（テスト）を行った。モデル検証を行った患者のうち、組織染色画像では腫瘍が抽出でき、さらに染色に成功した 13 症例で組織学的血管解析および視覚評価を行った。Radiomics 解析より T1、T2 それぞれで 8,175 個の特徴量を抽出し、次元削減後に Support vector machine 法を用いて脈管侵襲あり、なしを分類するモデルを構築した。組織免疫染色画像による腫瘍内血管特徴の定量的および視覚評価では、血管構造の組織形態を特異的および三次元的に抽出するために組織免疫染色を用いた。CUBIC（東京化成工業株式会社）を用いてがん組織の透明化を行い、血管内皮細胞とがん上皮細胞を免疫染色した。作製したサンプルは、共焦点レーザー顕微鏡（カールツァイス・LSM700）を使用し組織免疫染色画像撮影を行なった。視覚評価項目では、連続性や血管径、分岐点間の長さ、血管描出の鮮鋭度、画像全体に対する血管の分布を比較評価した。定量評価では、総分岐点数/血管面積、総血管長/血管面積、総終結点数/血管面積を評価した。

【結果】Radiomics 解析及び機械学習における脈管侵襲予測モデルの結果を表に示す。検証結果は、SMOTE なしの場合で精度、感度、特異度はそれぞれ 83.3%、51.8%、99.3%であり、平均 AUC は 91.6%であった。SMOTE ありの場合は精度、感度、特異度はそれぞれ 86.1%、83.2%、89.0%であり、平均 AUC は 92.1%であった。SMOTE なしの場合、感度と特異度で大きな差が生じていたが、SMOTE を使用することで感度と特異度の偏りは大きく改善し、さらに精度も改善した。脈管侵襲有りと無しの組織免疫染色画像で描出が良好なものに対して比較評価を行った。視覚評価では、脈管侵襲が有りの腫瘍内血管は脈管侵襲無しと比較して、血管描出が不鮮鋭であり、血管の分布が不均一な傾向を示した。脈管侵襲有りの血管特徴は、有意差は無いものの脈管侵襲無しと比較して血管面積あたりの分岐点数、終結点および血管長の値が大きくなった。この結果は、細く短い血管が高密度で細かく分岐し、不均一な分布で存在する傾向を示し、視覚評価の結果と一致した。以上の結果は、過去の文献にて示された腫瘍内血管新生の形態特徴と一致した傾向が見られた。子宮体癌脈管侵襲予測モデルでは、SMOTE を用いることにより予測精度はテストデータセットにおいて予測精度、AUC はそれぞれ 86.1%、0.92 となり、Sang らの研究で構築した予測モデルを 10%程度改善した。組織免疫染色画像評価より、脈管侵襲症例での均質性の向上は、腫瘍内血管新生の活発化により領域内血管が増加するものによることが示唆された。本研究では、SMOTE の使用により子宮体がんにおける脈管侵襲予測モデルの予測精度および感度を改善し、予測モデルの有用性が示唆された。さらに、免疫染色により得られた腫瘍内血管の傾向から、予測因子が示す均質性の指標が腫瘍内血管の特徴に起因することが示唆された。

脈管侵襲予測モデルテストデータの結果

	SMOTEなし	SMOTEあり
精度	83.3%	86.1%
感度	51.8%	83.2%
特異度	99.3%	89.0%
AUC	0.92	0.92

184. 機械学習による新規医療データ統合解析システムの構築

河原　大輔

広島大学 医系科学研究科 放射線腫瘍学

Key words：機械学習，Radiomics，組織免疫染色，SMOTE，血管新生

緒　言

子宮体がんの世界的な年齢標準化罹患率は 2020 年現在で 10 万人あたり 8.7 人であり、全世界の女性で 6 番目に多い癌である。そのうち死亡率は 10 万人あたりに 1.8 人となっている。日本では、以前の研究で、子宮体がんも 1985 年から 2015 年にかけて年齢標準化発生率が一貫して上昇していることが報告されている［1］。

子宮体がんは、子宮体部の内側にあり卵巣から分泌される卵胞ホルモンの作用をうけて月経をおこす子宮内膜という組織から発生し、子宮内膜がんとも呼ばれる。子宮体がんの生存率は、転移のない限局腫瘍の場合は 95.7%、領域リンパ節転移を有する症例では 73.2%、遠隔転移ありの症例では 20.1%まで低下する［2］。遠隔転移の経路として主にリンパ管を経由するリンパ行性転移と血管を経由する血行性転移の 2 種類がある。特に血管やリンパ管の中に癌細胞が入り込む状態を示す脈管侵襲は遠隔転移に大きく影響し、生存率の低下につながる。

子宮体がんの診断は、細胞・組織診等の病理検査や CT・MRI 等の医療画像検査により行われる。主な診断は、腫瘍の形質的な広がりを Computed Tomography（CT）や Magnetic Resonance Imaging（MRI）による画像診断で行われ、組織型や病理ステージは細胞・組織診にて決定される。このとき、脈管侵襲の評価については生検による組織診および摘出手術後の病理診断となり、治療前に撮影された MRI 画像による診断は困難である。

近年、治療前に撮影された多量の画像データ、臨床情報を活用して治療後の再発や転移等を予測する Radiomics 研究が行われている［3］。Radiomics 解析により画像から形状や不均一性など、多量の情報を定量化し、人工知能を活用することで予後との関連をモデル化することが可能である［4］。子宮体がんにおいては Radiomics 解析により収集した画像特徴量を使用した診断結果が放射線診断医による視覚評価に比べ改善したことが報告されており、さらに生存リスクを推定する研究や、子宮筋層浸潤を推定する研究などが行われ、臨床への活用の期待が高まっている［5］。しかし、予測因子となった Radiomics 特徴量がもつ生物学的特徴に関して解明することは行われていない。これを解明することで予後を改善する治療法の提案に貢献できる可能性があり、臨床的に大きな意義がある。

腫瘍微小環境など、生物的な情報を収集するため、組織免疫染色画像を用いて形態変化や細胞外マトリックス、腫瘍－免疫相互作用の解析などが行われ、腫瘍内の微小環境について可視化し把握できるようになった。しかし、腫瘍の微小環境は空間的に不均一であるにも拘わらず、腫瘍の一部分を切り抜き、2 次元画像撮影による視覚評価が行われてきた。近年では組織透明化技術の進歩により、腫瘍微小環境内の血管構造に関して 3 次元的に把握することも可能になった。

本研究の目的は、子宮体がんの脈管侵襲予測モデルの作成および検証を行い、対応する組織の組織染色画像より得られる血管内皮細胞および上皮細胞の形態から Radiomics 特徴量として表現される予測因子と組織学情報との関連を解析することで、脈管侵襲予測に寄与するがん微小環境構造内の特徴を解明することである。

方　法

1.　Radomics 解析及び機械学習による脈管侵襲予測モデル構築

子宮体がんの脈管侵襲予測モデル構築は、2014 年から 2020 年に広島大学病院にて子宮体がん摘出手術を受けた 153 症例を対象として行った。さらに、2021 年から 2022 年に広島大学病院にて子宮体がん摘出手術を受けた 32 症例を対象として予測モデルの検証（テスト）を行った。モデル検証を行った患者のうち、組織染色画像では腫瘍が抽出でき、さらに染色に成功した 13 症例で組織学的血管解析および視覚評価を行った。本研究の Radiomics 解析におけるワークフローを図 1 に示す。

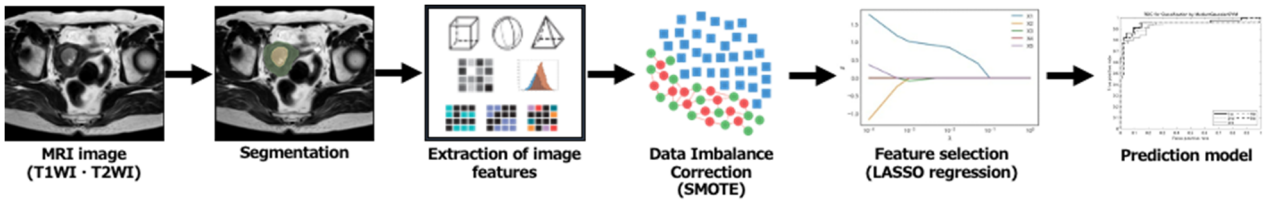


図 1.　Radiomics 解析のワークフロー
MRI 画像収集後、腫瘍領域の輪郭作成を行う Segmentation を行い、Radiomics 解析により画像特徴量抽出後、データ不均衡を補正する SMOTE を行い、特徴量を選別し予後予測モデルを構築する。

子宮体癌患者の術前 MRI（T1 強調画像 : T1WI、T2 強調画像 : T2WI）上で子宮体部を臨床標的体積（CTV）として描出し、次に示す関心領域（ROI）を作成した。CTV は子宮筋膜層、子宮体内部に存在する子宮筋腫を含み、腸管と重なる領域は除いた。CTV に対して図 2 に示す拡張、縮小した 4 つの ROI（CTV＋10 mm、CTV－10 mm、（CTV＋10 mm）－CTV、CTV－（CTV－10 mm））を作成した（図 2）。これらの ROI 内の画像特徴量を Radiomics 解析により抽出した。

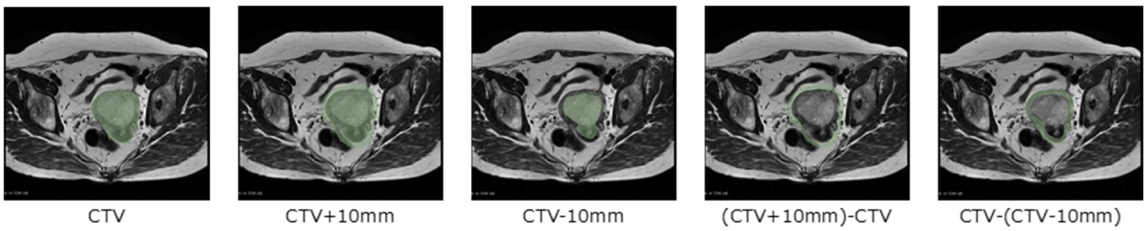


図 2.　Radiomics 特徴量抽出に用いた ROI
CTV は子宮体部全体の輪郭。CTV＋10 mm は CTV より 10 mm 拡大した領域、CTV－10 mm は CTV の周辺 10 mm を切り抜いた領域（CTV＋10 mm）－CTV は CTV の周辺 10 mm の領域、CTV－（CTV－10 mm）は CTV の内側の辺縁 10 mm の領域。

本研究の Radiomics 特徴量抽出では、元画像に加えて、元画像に対する 9 種類の Log フィルタと 8 種類のウェーブレットフィルタにて前処理を行い計 18 種類の画像を用いた。元画像に対する 15 個の形状特徴と、これら 18 種類の画像に対してそれぞれ 15 個の一次特徴、75 個のテクスチャ特徴を含む Radiomics 特徴量が抽出された。各 ROI における Radiomics 解析より T1、T2 それぞれで 8,175 個の特徴量を抽出した。次に、LASSO 回帰により特徴量選択を行い、機械学習（ML）分類法を用いて、子宮体がん患者を脈管侵襲有り（1）と無し（0）の 2 つのクラスに層別化した。機械学習は、Support vector machine 法を用いた。学習過程では、5 交差検証を使用し、さらにデータセットで不均衡が生じたため、オーバーサンプリング手法である Synthetic Minority Over-

sampling Technique（SMOTE）によりデータ不均衡補正を行なうことで、過学習を防いだ。予測性能は感度、特異度、精度、受信者動作特性（ROC）曲線下の面積（AUC）を用いて評価した。

2. 組織免疫染色画像による腫瘍内血管特徴の定量的および視覚評価

本研究において、血管構造の組織形態を特異的および三次元的に抽出するために組織免疫染色を用いた。組織透明化から画像処理および画像解析までの手順を図 3 に示す。CUBIC for animal（東京化成工業株式会社）を用いてがん組織の透明化を行った。免疫染色は、一次抗体は、CD31（PECAM-1）Mouse mAb と E-Cadherin Rabbit mAb を用いて、それぞれ血管内皮細胞（緑）とがん上皮細胞（赤）を免疫染色した。作製したサンプルは、共焦点レーザー顕微鏡（カールツァイス・LSM700）を使用し組織免疫染色画像撮影を行なった。撮影時、組織は Mounting solution for CUBIC-R 中に沈め、Tile：5×5、Slices：90、Range534.00 μm、Interval：6 を撮影条件とした。画像処理は、ZEN（bule edition）を用いて Orthogonal projection 処理を行い Maximum intensity projection（MIP）像の作製を行った。目視評価では、連続性や血管径、分岐点間の長さ、血管描出の鮮鋭度、画像全体に対する血管の分布を比較評価した。定量評価では、Image J 及び AngioTool を用いて血管解析を行った。解析項目は Total Number of Junctions／Vessels area（総分岐点数／血管面積）、Total Vessels Length／Vessels area（総血管長／血管面積）、Total Number of End Points／Vessels area（総終結点数／血管面積）とした。

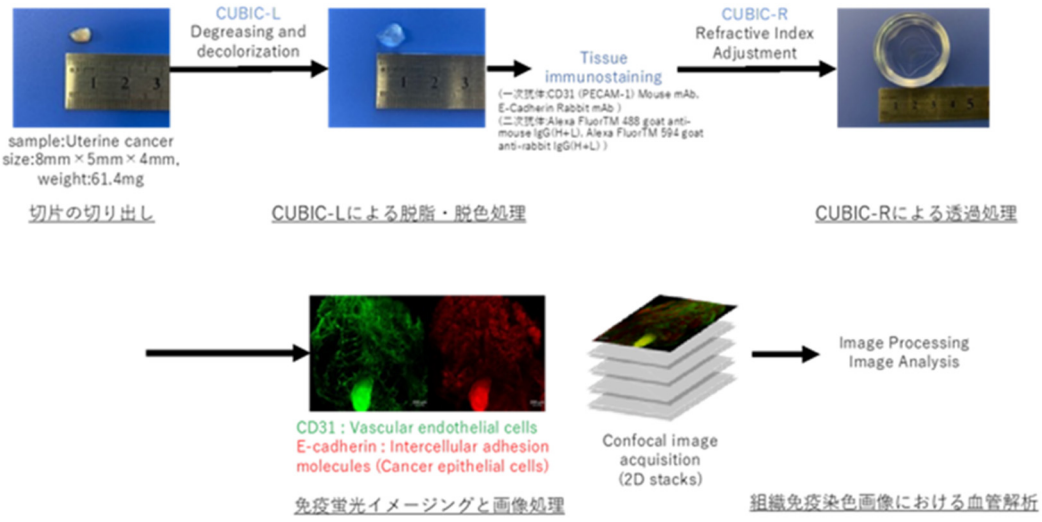


図 3. 組織透明化から画像処理および画像解析までの手順
組織切片を切り出し、CUBIC により組織透明化を行い、血管内皮細胞とがん上皮細胞を免疫染色した。画像をスタックし血管解析を行った。

結　果

1. Radomics 解析

脈管侵襲モデルの作成において、LASSO 回帰により 16 個（T1WI：7、T2WI：9）の特徴量が選択された。T1WI では CTV＋10 mm、（CTV＋10 mm）－CTV でグレーレベル強度値のばらつきおよび類似度、尖度に関する特徴量、CTV－10 mm で、小さい領域内での輝度値の分布やばらつき、結合分布の割合に関する特徴量が選択された。T2WI では、CTV で近接・近傍での輝度値の変化度やペアの発生頻度に対する特徴量、CTV－（CTV－10 mm）で、画像内の輝度値の均一性や一定領域における輝度値の均一性、非対称性、歪度に対する特徴量、（CTV＋10 mm）－CTV で、領域内グレーレベルの中央値や濃度の高低、画像内の輝度値の均一性と歪度に対

する特徴量が選択された。また、特徴量の定義より、総じて脈管侵襲有りのグループが均質性は高く、テクスチャが細かく、グレーレベルおよび領域内の輝度値のばらつきが比較的低いことが示された。

事前に作成した脈管侵襲予測モデルの予測性能をテストデータで検証した予測結果を表 1 に示す。検証結果は、SMOTE なしの場合で精度、感度、特異度はそれぞれ 83.3%、51.8%、99.3%であり、平均 AUC は 91.6%であった。SMOTE ありの場合は精度、感度、特異度はそれぞれ 86.1%、83.2%、89.0%であり、平均 AUC は 92.1%であった。SMOTE なしの場合、感度と特異度で大きな差が生じていたが、SMOTE を使用することで感度と特異度の偏りは大きく改善し、さらに精度も改善した。

表 1. 脈管侵襲予測モデルテストデータの結果

	SMOTEなし	SMOTEあり
精度	83.3%	86.1%
感度	51.8%	83.2%
特異度	99.3%	89.0%
AUC	0.92	0.92

SMOTEありの方がデータ不均衡を補正ており精度が改善している。予測精度も感度と特異度の偏りが低減した。

2. 組織免疫染色による腫瘍内血管特徴の視覚評価及び血管解析

脈管侵襲有りと無しの組織免疫染色画像で描出が良好なものに対して比較評価を行った。脈管侵襲の有無それぞれの組織免疫染色画像を図 4 に示す。視覚評価では、脈管侵襲が有りの腫瘍内血管は脈管侵襲無しと比較して、血管描出が不鮮鋭であり、血管の分布が不均一な傾向を示した。AngioTool による血管解析の結果を表 2 に示す。脈管侵襲有りの血管特徴は、有意差は無いものの脈管侵襲無しと比較して血管面積あたりの分岐点数、終結点および血管長の値が大きくなった。この結果は、細く短い血管が高密度で細かく分岐し、不均一な分布で存在する傾向を示し、視覚評価の結果と一致した。以上の結果は、過去の文献にて示された腫瘍内血管新生の形態特徴と一致した傾向が見られた。

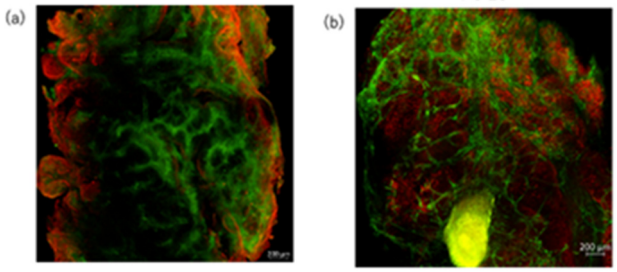


図 4. 脈管侵襲有り（a）と無し（b）の組織免疫染色画像
脈管侵襲が有りの場合は腫瘍内血管描出が不鮮鋭であり、血管の分布が不均一であった。一方で脈管侵襲が無しの腫瘍内血管描出は鮮明であった。

表 2. AngioTool での解析結果

	Average ± SD		p-value
	Yes	No	
Total Number of Junctions/Vessels area	0.00074 ± 0.00017	0.00063 ± 0.00012	0.18
Total Vessels Length/Vessels area	0.062 ± 0.0041	0.058 ± 0.0039	0.14
Total Number of End Points/Vessels area	0.00089 ± 0.00017	0.00078 ± 0.00017	0.26

Yesは脈管侵襲有、Noは脈管侵襲無しの症例群を示す。T検定による有意差検定を行ったが有意差は認められなかった。症例数が少数であることが要因と考えており、脈管侵襲有りの血管特徴は、有意差は無いものの脈管侵襲無しと比較して血管面積あたりの分岐点数、終結点および血管長の値が大きい傾向を示している。

考　察

子宮体癌における脈管侵襲、過去の研究において予後不良因子であることが報告されおり、さらに脈管侵襲の有無は診断において重要なリスク因子であるが、術後の病理検査結果まで LVSI を判定できない。これを術前の医療画像から予測することができれば、手術手技や経過観察間隔などの治療方針を決定する上で有用な情報となる。Sang らの研究では、腫瘍径や腫瘍グレード、子宮筋層浸潤の深さなどの子宮体癌の臨床病理学的因子から脈管侵襲指数を作成し、脈管侵襲の予測精度は AUC が 0.839 であった［6］。

本研究では、手術前に撮影された MRI 画像（T1WI・T2WI）に対して Radiomics 解析及び機械学習を用いて脈管侵襲の有無の予測精度改善を図った。テストデータセットにおいて予測精度、AUC はそれぞれ 92.7%、0.936 となり、Sang らの研究で構築した予測モデルを 10%程度改善し高率な脈管侵襲予測モデルであった。選択された特徴量より、脈管侵襲無しと比較して脈管侵襲有りの方が腫瘍部分を多く含む CTV 内領域において、T2WI で低い信号値かつばらつきが小さい特徴を示した。これは血液、水分の割合が低い状態を示しており、腫瘍内に十分に血液や栄養が送られていない特徴を示す。過去の文献では、子宮平滑筋腫症で生じる脈管侵襲有りの MRI 画像所見として、均質な低 T2 信号強度を示したことが報告されており、本研究で選択された画像特徴量と一致した［7］。また、T1WI より CTV 内の画像特徴として、グレーレベル値の小さいサイズゾーンの結合分布の割合が大きいことや均質性が高いこと、テクスチャが細かいことを示した。これは血管新生により新たに構築された血管ネットワークを画像上で検出している可能性がある。組織免疫染色画像における視覚評価では、脈管侵襲有りの腫瘍内血管での血管径の不均一性や不規則な分岐、領域全体に対する血管分布の不均一性が観察された。これは、過去の文献で示されているがん微小環境における血管新生の特徴を示すものと考えられる［8］。さらに、AngioTool による血管解析結果より、脈管侵襲有りで血管新生の不規則な分岐や血管径のばらつき、領域内分布の不均一な血管分布などの特徴を表現する結果となり、視覚評価と同様の結果を示した［9］。 医療画像における Radiomics 解析、手術により摘出した検体における組織免疫染色画像評価より、脈管侵襲症例では腫瘍内の血液や水分の割合が低下するため、腫瘍内血管新生の活発化により領域内血管が増加していることが考えられる［10］。医療画像解析結果と組織免疫染色画像解析結果で一貫した結果が得られたことから、これまでブラックボックスであった医療画像の生物学的な解釈が可能であることを示した。

本研究において、いくつかの課題がある。1 つ目に、子宮体がんにおける脈管侵襲の予測能力と頑健性を向上させるために、新規症例での検証および予測因子の意味づけを行った。しかし、新規症例の医療画像は単一の医療機関ではなく複数の医療機関にて取得された医療画像であった。従って、医療画像の撮影条件や撮影機器の違いにより、画質が異なることで予測精度へ影響する可能性がある。2 つ目に、組織学的特徴を算出したがん組織検体は、がん組織全体からごく一部の領域の特徴であり、がん組織全体の血管分布および血管量を再現していない可能性がある。がん組織の固有特徴を完全に反映するにはがん組織全体および周辺組織を含めた血管描出を行う必要がある。そのためには、がん組織全体の提供と大型サンプルから血管およびその他の組織・細胞を撮影す

ることができる大規模な研究設備が必要である。もしくは、1 つのがん組織から位置情報を付与した複数の組織を観察し、がん組織全体を表現する必要がある。3 つ目に、組織免疫染色および共焦点顕微鏡撮影を行う中でノイズが血管描出精度へ影響することが考えられる。今後は組織免疫染色の再現性に関して同じ症例を複数に分割して染色することで検証し、さらに共焦点顕微鏡撮影を行う上でのバックグラウンド補正を行うことを検討している。

共同研究者・謝辞

本研究の共同研究者は、統合生命科学研究科統合生命科学専攻基礎生物プログラム発生生物学研究室の菊池裕教授、高橋治子助教、広島大学病院産婦人科の古宇家正先生、関根仁樹先生である。

本研究を遂行するにあたり、幅広い知識と経験のもと興味深いテーマやアイデアを与えてくださり、基礎生物学研究についてのご指導をいただきました広島大学大学院統合生命科学研究科統合生命科学専攻基礎生物プログラム発生生物学研究室の菊池裕教授、高橋治子助教に深く感謝いたしますとともに、御礼申し上げます。

文　献

1） Lortet‐Tieulent J, Ferlay J, Bray F, Jemal A. International patterns and trends in endometrial cancer incidence, 1978‐2013. J Natl Cancer Inst. 2018;110(4):354‐361. PMID: 29045681.

2） Ito Y, Miyashiro I, Ito H, Hosono S, Chihara D, Nakata-Yamada K, Nakayama M, Matsuzaka M, Hattori M, Sugiyama H, Oze I, Tanaka R, Nomura E, Nishino Y, Matsuda T, Ioka A, Tsukuma H, Nakayama T; J-CANSIS Research Group. Long-term survival and conditional survival of cancer patients in Japan using population-based cancer registry data. PMID: 25183551.

3） Cui L, Yu T, Kan Y, Dong Y, Luo Y, Jiang X. Multi-parametric MRI-based peritumoral radiomics on prediction of lymph-vascular space invasion in early-stage cervical cancer. Diagn Interv Radiol. 2022 Jul;28(4):312-321. PMID: 35731710.

4） Li Y, Zhang Y, Fang Q, Zhang X, Hou P, Wu H, Wang X. Radiomics analysis of [18F]FDG PET/CT for microvascular invasion and prognosis prediction in very-early- and early-stage hepatocellular carcinoma. Eur J Nucl Med Mol Imaging. 2021 Jul;48(8):2599-2614. PMID: 34169367.

5） Xu X, Li H, Wang S, et al. Multiplanar MRI-based predictive model for preoperative assessment of lymph node metastasis in endometrial cancer. Front Oncol 2019; 9: 1–11. PMID: 31649877.

6） Kim SI, Yoon JH, Lee SJ, Song MJ, Kim JH, Lee HN, Jung G, Yoo JG. Prediction of lymphovascular space invasion in patients with endometrial cancer. Int J Med Sci. 2021 Jun 1;18(13):2828-2834. PMID: 34220310.

7） LeGout, Ryan E. Bailey, Candice W. Bolan, Andrew W. Bowman, Frank Chen, Joseph G. Cernigliaro, and Lauren F. Alexander. Multimodality Imaging of Abdominopelvic Tumors with Venous Invasion Jordan D. RadioGraphics 2020 40:7, 2098-2116. PMID: 33064623.

8） Tien YW, Chang KJ, Jeng YM, Lee PH, Wu MS, Lin JT, Hsu SM. Tumor angiogenesis and its possible role in intravasation of colorectal epithelial cells. Clin Cancer Res. 2001 Jun;7(6):1627-32. PMID: 11410499. PMID: 11410499.

9） McDonald DM, Choyke PL. Imaging of angiogenesis: from microscope to clinic. Nat Med. 2003 Jun;9(6):713-25. PMID: 12778170.

10) Baghban R, Roshangar L, Jahanban-Esfahlan R, Seidi K, Ebrahimi-Kalan A, Jaymand M, Kolahian S, Javaheri T, Zare P. Tumor microenvironment complexity and therapeutic implications at a glance. Cell Commun Signal. 2020 Apr 7;18(1):59. PMID: 32264958