

学位論文 (博士)

Exploration and time-serial validation of logistic regression models composed of multiple laboratory tests for early detection of HCV-associated hepatocellular carcinoma

(C型肝炎関連肝細胞癌の早期発見を目的とした臨床検査による最適ロジスティック診断モデルの探索と遡り検証からみた臨床的有用性)

氏名 小林 利彦
所属 山口大学大学院医学系研究科
医学専攻 医療情報判断学講座

令和3年6月

目次

<u>1. 要旨</u>	<u>2</u>
<u>2. 緒言.....</u>	<u>3</u>
<u>3. 目的.....</u>	<u>4</u>
<u>4. 方法.....</u>	<u>4</u>
(1) 対象	4
(2) 方法	6
(3) 解析	7
(4) 倫理的配慮.....	8
<u>5. 結果.....</u>	<u>8</u>
<u>6. 考察.....</u>	<u>17</u>
<u>7. 結語.....</u>	<u>18</u>
<u>8. 謝辞.....</u>	<u>18</u>
<u>9. 参考文献.....</u>	<u>19</u>

1. 要旨

C型慢性肝炎・肝硬変から発症する肝細胞癌（HCC）の定期サーベイランスにおける腫瘍マーカーを含む日常臨床検査、およびその時系列変化の臨床的有用性を明らかにするため、長期の経過観察中にHCCを発症した129例（HCC群）と未発症例323例（非HCC群）を対象に日常臨床検査による診断モデルの開発とその時系列変化を検討した。

対象症例から年齢、性別とFIB-4 indexを変数とした傾向スコアマッチングにより抽出された各102例を解析対象とし、AFP、PIVKA-II、および日常臨床検査項目について、その経時的変化を基準日より3ヶ月毎に1年間遡って検討した。2群間で有意であった検査項目を用いたHCC診断のロジスティック回帰モデル（LRM）を作成後、遡り期間の検査値にモデルを適用し得られたHCC予測値の経時的変化からHCCの早期診断への寄与について検討した。

AFP、AFP/(AST×ALT)、ASTを説明変数としたLRMが最適であり、式は以下の通りであった。

$$\text{肝細胞癌確率} = 1 / (1 + \exp(-19.514 + 22.084 \times \text{AFP}^{-0.3} - 1.698 \times [\text{AFP}/(\text{AST} \times \text{ALT})]^{-0.3} + 4.238 \times \ln(\text{AST})))$$

このLRMでの1年前、6ヶ月前、3ヶ月前、診断時におけるROC曲線下面積（AUC）は、0.654、0.786、0.834、0.849と1年前から診断時に向けて経時的な上昇を認め、診断時のAUCは6ヶ月、9ヶ月、1年前のAUCと比べ、また、3ヶ月前のAUCは9ヶ月、1年前のAUCと比べ有意に高値であった。このLRMで偽陽性率が20%となるHCC予測値を閾値としそれを超える場合を陽性とした時、1年前から診断時にかけて18%の陽性率の単調かつ漸増的な上昇変化を認めた。この陽転時点において詳細な画像検査を実施することで数ヶ月の早期発見に寄与できる可能性が示唆された。

2. 緒言

肝細胞癌（HCC）は頻度が高い悪性腫瘍として知られ、日本における年間死亡者数は約 2 万 5 千人と肺癌、大腸癌、胃癌、膵臓癌に次いで第 5 位となっている¹⁾。1990 年代には hepatitis C virus（HCV）がその原因の 6~7 割を占めていたが、2015 年には非 B 非 C によるものが 3 割を超え著しい増加傾向にある²⁾。さらに近年の HCV 感染症に対する直接作用型抗ウイルス療法の進歩により大多数の患者が持続的なウイルス学的著効（SVR）を達成し、これらの患者での HCC の発生率は減少している。しかし、今だ HCV の持続感染による慢性肝炎あるいは肝硬変の患者からの HCC 発症が全体の 5 割程度を占める状況にある。また、特に肝線維化の進行した患者では、SVR 後でも HCC を発症するリスクがあり³⁻⁴⁾、例えば、Kanwal らは、肝硬変患者で年間 1.38% の HCC 発症率を報告している³⁾。そのため、HCC の早期診断は未だに重要な課題として残されており⁵⁾、HCC の発癌リスクが高い C 型慢性肝炎・肝硬変患者に対する定期的なサーベイランス検査を少なくとも 6 ヶ月間隔で実施することが推奨されている⁶⁾。

本来、このようなサーベイランス検査の有用性の判定は、非サーベイランス群とのランダム化比較試験により見いだされた HCC における治療予後がどの程度改善するかによって評価される⁷⁻⁸⁾。しかし、現実的には倫理的な配慮から実施困難であり、横断的研究による検査等や診断モデルにおける診断能の優劣から評価されている⁹⁻¹¹⁾。

一般的に HCC の早期診断に用いられているサーベイランス検査は α -fetoprotein（AFP）や Protein Induced by Vitamin K Absence-II（PIVKA-II）などの腫瘍マーカーおよび腹部超音波（US）、CT、MRI などの画像検査である。また近年では、ソナゾイド造影剤を使用した US（CEUS）や¹²⁻¹³⁾、Gd-EOB-DTPA 造影剤を使用した MRI と言った診断性能のより高い画像検査が導入されている¹⁴⁾。

日本においては、画像検査による定期的なサーベイランスは通常、非造影 US または造影 CT によって行われ、これらの画像検査で HCC が疑われる場合は、それらの造影 US や MRI、血管造影等の追加検査で詳しく精査されている。

一方、HCC の診断に対する腫瘍マーカー等の臨床検査の診断特性はこれまでも多く報告されているが、これまでのシステマティックレビュー等の報告では AFP および PIVKA-II の感度は十分なものではないとされている¹⁵⁻¹⁶⁾。特に Barcelona Clinic Liver Cancer (BCLC) 病期分類システムにおいて、早期ステージ（Stage 0 & A）に該当する腫瘍径 3cm 未満、腫瘍個数 3 個以下に限った HCC では、診断能はさらに低く、これらの腫瘍マーカーの有用性は限定的とされている¹⁷⁾。しかし、これらの研究報告は横断的研究によるものであり、経時的な変化を系統的に検討した研究は少なく、その評価は未だ定まっていない¹⁸⁻¹⁹⁾。

3. 目的

そこで、今回、C 型慢性肝炎・肝硬変にて長期の経過観察中に肝細胞癌を発症した症例と未発症例を対象に、腫瘍マーカーおよび日常臨床で頻用される臨床検査データを用いて、最適ロジスティック診断モデルの探索と、その経時的な変化を遡って検証することで、肝細胞癌の早期診断に寄与できるかを明らかにすることを目的に、検討を行ったので報告する。

4. 方法

(1) 対象

1) 対象者

1994 年 1 月 1 日～2018 年 9 月 30 日までの期間に山口大学医学部附属病院の肝疾患専門外来で C 型慢性肝炎および肝硬変と診断された患者の中から、i) HCV-RNA または HCV 抗体が陽性、ii) 過去に肝細胞癌の診断および治療歴がない、iii) 経過観察のために 1 年以上の通院歴がある、iv) AFP あるいは PIVKA-II が年 3 回以上測定されている、これら全てを満たす症例を対象とした。このうち、HBsAg、HBV-DNA、HIVAg/Ab のいずれかが陽性、自己免疫性肝炎、原発性胆汁性肝硬変、アルコール性肝炎・肝硬変の診断例は除外した。その結果、452 例が対象となり、その内訳は、経過中に肝細胞癌を発症した群（HCC 群）129 症例、発症していない群（非 HCC 群）323 症例であった（図 1）。

HCC の診断は、主に画像検査（造影 US、造影 CT、造影 MRI、血管造影）または肝生検による組織病理学的検査での特徴的な所見に基づいて行われた。当院の肝疾患専門外来では、通常、慢性肝炎の場合、造影なしの US（US-NE）が 6 か月ごとに実施され、肝硬変の場合には、非 SVR の場合には 3～4 か月ごとに、SVR の場合では 6 ヶ月ごとに US-NE または造影 CT のいずれかの検査が実施されている。また、それらの画像検査で HCC が疑われる患者には、さらに CEUS（2007 年からはソナゾイドを使用）または造影 MRI（2008 年から Gd-EOB-DTPA を使用）が実施されている。

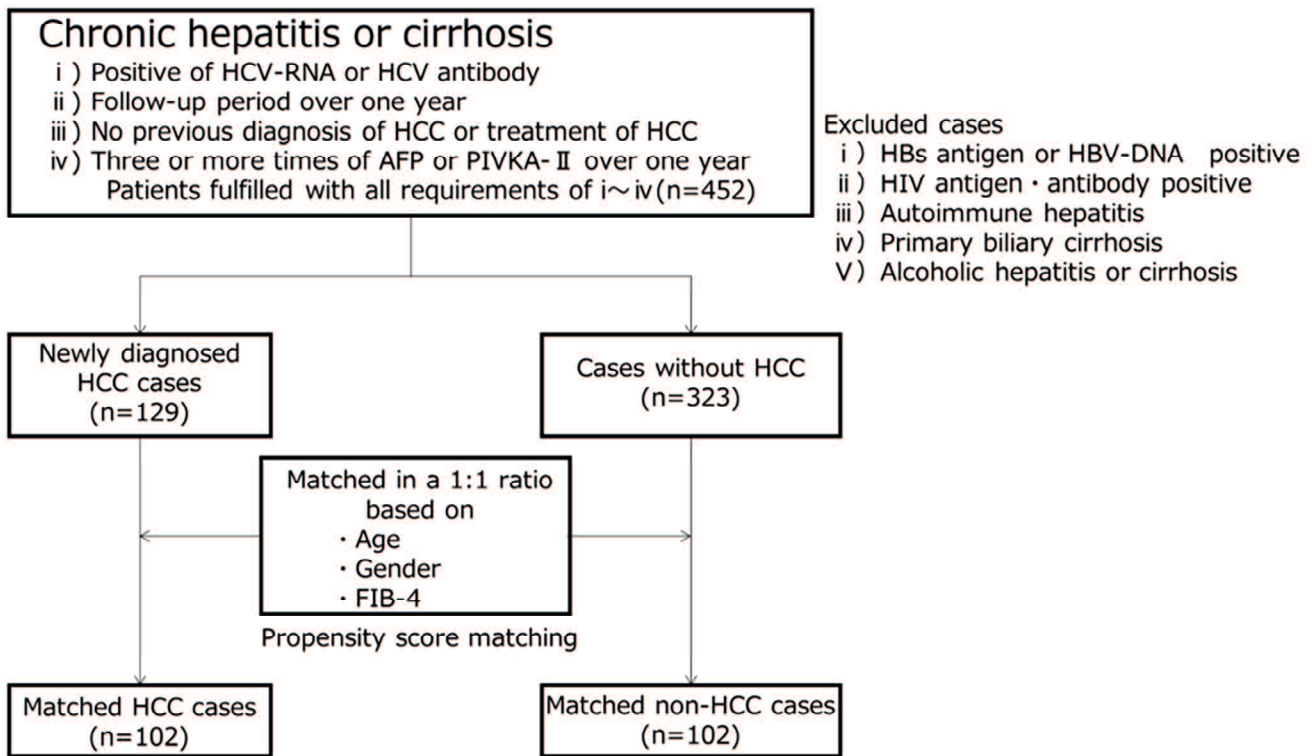


図 1. 症例選択のフロー

2) 解析対象項目

解析には年齢と性別に加え、検査項目では AFP, PIVKA-II, total protein (TP), albumin (ALB), aspartate aminotransferase (AST), alanine aminotransferase (ALT), lactate dehydrogenase (LD), alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), cholinesterase (ChE), creatinine (CRE), urea nitrogen (UN), ammonia (NH₃), glucose (GLU), sodium (Na), potassium (K), chloride (Cl), total bilirubin (TBIL), direct bilirubin (DBIL), White blood cell (WBC), hemoglobin (Hb), platelet (PLT), activated partial thromboplastin time (APTT), prothrombin time (PT)、また、肝細胞癌診断の指標として報告されている AFP/(AST×ALT)²⁰⁾、さらに肝線維化の指標として多用される FIB-4 index (FIB-4)²¹⁾を加え、計 26 項目を解析対象とした。FIB-4 は、下記の計算式により算出した。

$$\text{FIB-4} = (\text{年齢} \times \text{AST}) / (\text{PLT} \times \text{ALT}^{0.5})$$

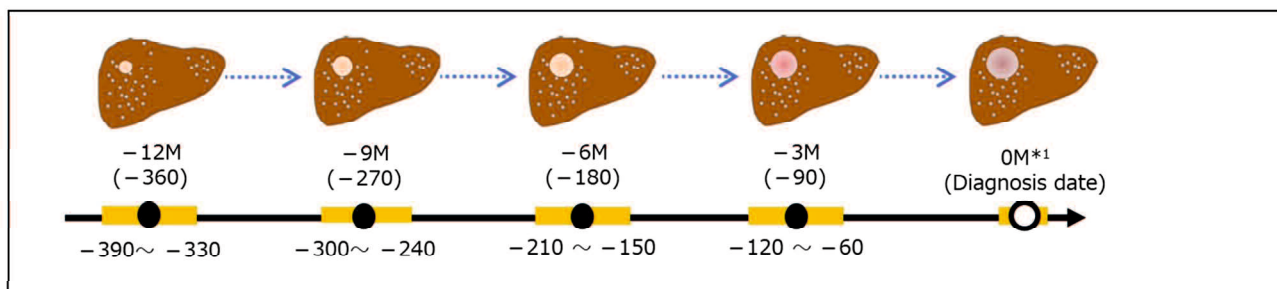
なお、ワーファリン服用中の患者では、服用時における PT と PIVKA-II の検査結果は除外した。

(2) 方法

1) 遡り期間の検査値データの取得

HCC 群においては、肝細胞癌診断時を基準日として 3 ヶ月単位で 1 年前まで過去に遡り各検査データを取得した。すなわち、診断時を 0 日として、+1~30 日の検査値（複数回の検査値がある場合は、その平均）を診断時のデータ（0M）、-60~-120 日の検査値を 3 ヶ月前のデータ（-3M）、-150~-210 日の検査値を 6 ヶ月前のデータ（-6M）、-240~-300 日の検査値を 9 ヶ月前のデータ（-9M）、-330~-390 日の検査値を 12 ヶ月前のデータ（-12M）とした。非 HCC 群においては、直近の検査日に近接した期間での肝細胞癌診断の可能性を考慮して最終の血液検査日より 1 年以前に AFP の検査が実施された日を基準日（0M）として、HCC 群と同様に基準日から過去に遡ってデータを取得した（図 2）。

Data acquisition scheme for HCC cases



Data acquisition scheme for non-HCC cases

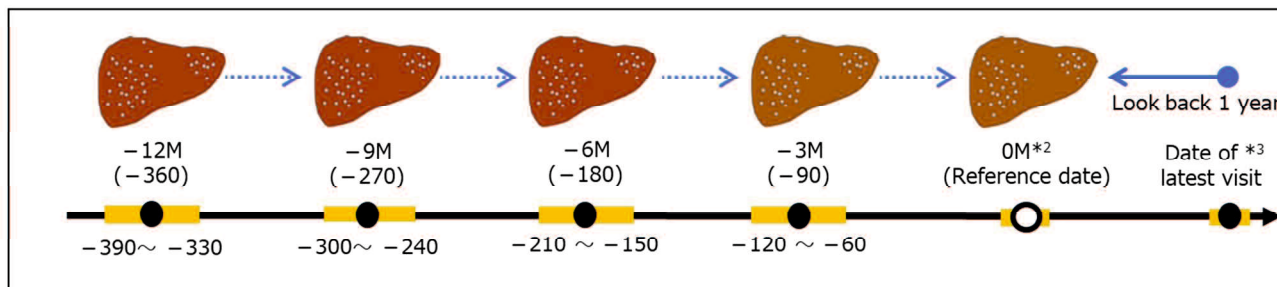


図 2. HCC 群および非 HCC 群のデータ取得方法

HCC 群では、診断時を 0M（*1）とし、3 ヶ月間隔（±30 日）で最長 1 年間遡ってデータを取得した。非 HCC 群については、診断の正確性を確保するため、最新の受診日（*3）の 1 年前を基準日 0M（*2）と設定した。

2) 検査値データの長期連続性の保持

長期間に渡る対象患者の収集であったため、期間中に検査の単位・測定法・試薬等の変更がある場合には、その必要性によって標準主軸回帰式にて得られた変更前後の測定値の変換式を用いて、過去の方法の検査値を最新の検査法に合わせ検査値の換算を行った。期間中に変換が必要であった項目と回数は AFP、PIVKA-II、NH₃ で 2 回、ChE、WBC、APTT、PT で 1 回であった。

3) 症例の傾向スコアマッチング

対象で得られた HCC 群：129 例と非 HCC 群：323 例の中で年齢、性別、FIB-4 によるロジスティック回帰モデルから得られた傾向スコアによって 1:1 のマッチングを行った。その際、Caliper は、 $0.25 \times SD$ 以下が推奨されており、最終的には Caliper を $0.1 \times SD$ とした²²⁾。その結果、マッチング後の対象症例は各々102例となった。

(3) 解析

1) 単変量統計解析

マッチング症例における HCC 群と非 HCC 群の 2 群でのそれぞれの基準日 (0M) での各検査結果について基本統計量を求め、頻度は χ^2 、あるいは Fisher の直接確率検定、年齢および検査項目における 2 群間差の検定は Mann-Whitney U test で行った。また、全ての解析において有意水準は $P = 0.05$ とした。

HCC 群、非 HCC 群の 2 群で有意差を認めた項目において、群別での全症例の検査項目における -12M、-9M、-6M、-3M、0M の経時的変化について Jonckheere-Terpstra trend test による傾向検定を実施した。

なお、本研究での統計解析には Stata 16 (StataCorp LLC, Texas USA) を用いた。

2) 肝細胞癌診断のロジスティック回帰 (LR) モデルの構築

HCC 群 (0M) と非 HCC 群 (0M) のマッチング症例において HCC の有無を目的変数、単変量解析において 2 群間に有意差を認めた検査項目を説明変数として組み入れ、探索的に診断能の高い LR モデル作成の検討を行った。この際、全症例の半数をモデル構築のためのトレーニングデータセットと残りの症例を検証データセットとして無作為に分類しモデルの検証を行った。また、2 つのデータセット間の AUC モデルに組み入れる説明変数の項目数を HCC 群の症例数の 1/10 を上限として変数選択を行い²³⁾、モデルの選択には赤池の情報量基準 (AIC) を参考にした。偏った分布の検査項目は、対数変換、べき乗変換を使用して正規分布となるようにした。またモデルには変換した値を用いた。次に、0M の LR モデルを各遡り期間 (HCC 群と非 HCC 群の -3M、-6M、-9M、-12M) のデータに当てはめ、それぞれの遡り期間における ROC 曲線下面積 (AUC) を求め、診断性能の指標とした。そして、各遡り期間の AUC の有意差検定を行い、それぞれの期間での AUC の差異を比較した。

3) 遡り期間での陽性率の比較

LR モデルで得られた 0M における ROC 曲線から特異度 80%となる肝細胞癌予測値をカットオフ値として設定し、各遡り期間での欠損値がない症例における陽性率（カットオフ値を超える症例の割合）を求めた。同様に AFP 単独および PIVKA-II 単独での特異度 80%における陽性率も比較した。この陽性率の経時的な変化を見ることにより、どの程度 HCC の早期診断が可能となり得るかを検討した。

サブ解析として、BCLC 病期分類システムでの早期ステージ（Stage 0 & A）の HCC 症例に限定した場合の、陽性率を比較した。

(4) 倫理的配慮

本研究は、当院での倫理審査委員会（IRB）の承認（H30-127）を得るとともに、既存の診療情報を活用したことから、患者同意はオプトアウトとし、参加拒否の機会を公知した。

5. 結果

1) 対象患者の基本的特性

表 1 にマッチング症例での 0M における HCC 群と非 HCC 群の各項目の基本統計量の結果を示した。AFP、PIVKA-II、AST、ALT、GGT、SVR で $P < 0.001$ 、ALP で $P < 0.01$ 、ALB、K、AFP/(AST×ALT)で $P < 0.05$ の 2 群間での有意差を認めた。

統計的に有意であった項目は LR モデルの説明変数として用いた。また、AFP、AFP/(AST×ALT)は -0.3 乗で変換、PIVKA-II、AST、ALT、GGT、ALP は対数変換し用いた。

HCC 群における BCLC 病期分類システムでの、早期ステージ（Stage 0 & A）は全体の 74%であった。

表 1. HCC 群と非 HCC 群でのそれぞれの基本的特性結果

Variables (units)	HCC		non-HCC		P-value
	n	Median (25–75%)	n	Median (25–75%)	
Age (years)	102	72 (66–77)	102	71 (67–76)	0.761
Male (%)	102	50.0	102	43.1	0.413
AFP (µg/L)	92	16.6 (8.2–99.7)	102	4.7 (3.0–9.3)	***0.000
PIVKA-II (mAU/mL)	86	30.5 (17.9–69.5)	61	19.4 (14.1–30.3)	***0.000
TP (g/L)	100	74 (70–79)	93	74 (70–77)	0.312
ALB (g/L)	101	37 (33–41)	102	40 (33–43)	*0.030
AST (U/L)	102	48 (35–71)	102	32 (25–51)	***0.000
ALT (U/L)	102	40 (26–63)	102	22 (14–35)	***0.000
LDH (U/L)	101	220 (189–265)	94	215 (192–263)	0.965
ALP (U/L)	102	307 (250–414)	95	265 (205–353)	**0.006
GGT (U/L)	102	40 (26–57)	101	22 (16–37)	***0.000
ChE (U/L)	95	187 (132–267)	58	175 (118–277)	0.712
CRE (µmol/L)	101	69.0 (57.5–83.1)	101	68.1 (58.3–85.7)	0.565
UN (mmol/L)	101	5.6 (4.6–7.1)	101	5.7 (4.7–7.1)	0.615
NH3 (µmol/L)	52	31.6 (26.1–43.9)	28	23.7 (18.0–45.8)	0.117
GLU (mmol/L)	89	5.7 (5.3–6.4)	93	5.7 (5.3–6.3)	0.990
Na (mmol/L)	95	141 (139–142)	90	140 (139–142)	0.498
K (mmol/L)	95	4.1 (3.9–4.4)	90	4.3 (4.0–4.5)	*0.023
Cl (mmol/L)	95	106 (105–108)	90	106 (104–108)	0.985
TBIL (µmol/L)	102	13.7 (12.0–20.5)	102	13.7 (10.3–20.5)	0.786
DBIL (µmol/L)	89	3.4 (3.4–6.8)	50	5.1 (3.4–6.8)	0.083
WBC (10 ⁶ /L)	102	4155 (3585–5750)	102	4076 (3542–4840)	0.183
Hb (g/L)	102	128 (115–140)	102	125 (110–137)	0.134
PLT (10 ¹⁰ /L)	102	11.6 (8.3–14.9)	102	11.9 (8.6–14.3)	0.743
APTT (sec)	71	34.6 (31.2–37.4)	11	36.9 (32.2–38.9)	0.298
PT (%)	76	85.0 (75.2–92.2)	46	79.5 (62.7–93.8)	0.251
AFP/(AST×ALT)	92	0.010 (0.004–0.045)	102	0.008 (0.005–0.012)	*0.047
FIB-4 index	102	5.0 (3.3–7.4)	102	4.7 (3.2–6.5)	0.445
SVR (%)	100	12.0	102	53.9	***0.000
Tumor diameter (cm)	99	2.1 (1.5–2.8)			
Number of tumors (%)	99	single (71) two (18) three or more (11)			
BCLC staging system Stage 0 & A (%)	102	74			

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$

マッチング症例における HCC 群と非 HCC 群での各検査項目の基本的特性と P-value を示す。

(* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$)

2) 経時的な結果値の上昇と傾向検定

マッチング症例における HCC 群と非 HCC 群での傾向検定の結果、非 HCC 群では全ての項目で有意な傾向を認めなかったが、HCC 群では AFP ($P = 0.016$)、PIVKA-II ($P = 0.000$)、AFP/(AST×ALT) ($P = 0.016$) の 3 項目で有意な上昇傾向を認めた。これら 3 項目の遡り期間における経時の上昇を図 3 に示す。

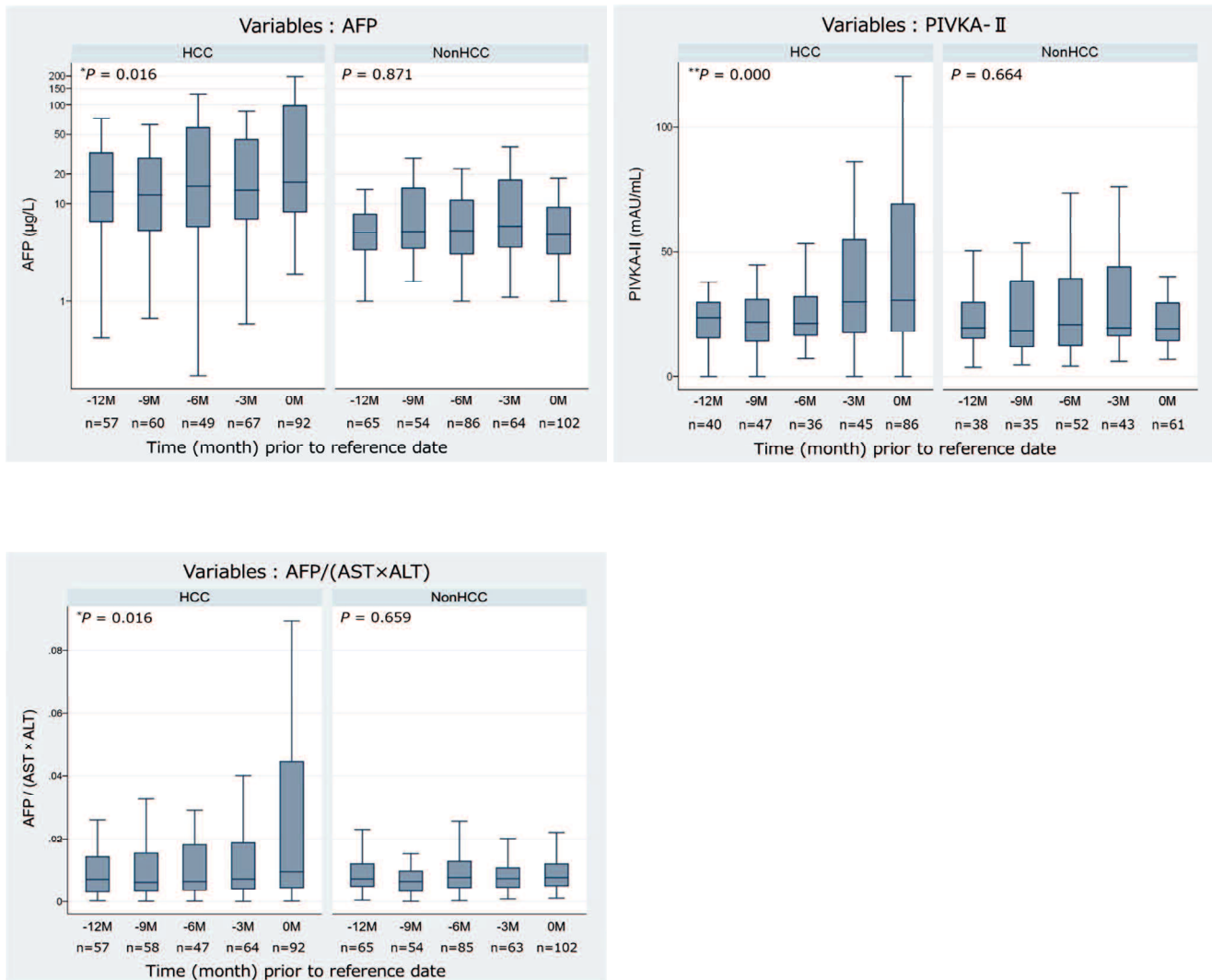


図 3. HCC 群と非 HCC 群における AFP、PIVKA-II、AFP/(AST×ALT) の時系列変化

3) 腫瘍マーカーおよび LR モデル (LRM) の結果

AFP、PIVKA-II、AFP/(AST×ALT)および、HCC 群と非 HCC 群の 2 群間 (0M) で有意差を認めたその他の項目を説明変数として探索的に LR 解析を実施した結果 (LRM1~LRM5) を表 2 に示す。

LRM1 の説明変数は、AFP、PIVKA-II、LRM2 では AFP、PIVKA-II、AFP/(AST×ALT)であり、残りのモデル (LRM3~LRM5) は、腫瘍マーカーを含むさまざまな変数の組み合わせで構築したモデルである。LRM 1、2、3、4、5 の AUC は、それぞれ 0.787、0.788、0.845、0.869、0.830 であった。最も AUC が高かったのは LRM4 の 0.869 であった。

トレーニングデータセットで構築した各 LRM の信頼性を評価するために、検証データセットに適用した結果、どの LRM においても、トレーニングデータセットと同等の AUC であり、2 つのデータセット間の AUC に有意差は認めなかった。

これらの結果から、今後の分析は、AUC が最も高く AIC が低い LRM3 および LRM4 を選択した。比較として、AFP 単独と PIVKA-II 単独の診断精度も同様に評価した。サブ解析として、SVR の影響の可能性を考慮して、LRM4 に SVR を追加して調べた。その結果、SVR は有意ではなく ($P = 0.290$)、LRM4 の診断性能は SVR の有無にかかわらず変化しなかった (AUC: 0.874)。データセット間に有意差はなかったことと、データ数確保のため、今後の分析では、LRM3 と LRM4 は、トレーニングデータセットと検証データセットを組み合わせたデータセットを使用して再構築した。その結果、LRM3 と LRM4 の AUC は、それぞれ 0.850 と 0.862 であった。AFP 単独と PIVKA-II 単独の AUC はそれぞれ 0.810 と 0.660 と推計された。

図 4 に示すように、LRM3 と LRM4 の AUC (0.850 と 0.862) は、AFP 単独 (0.810) および PIVKA-II 単独 (0.610) の AUC よりも有意に高く診断性能が良い結果であった。一方、LRM3 と LRM4 においては AUC に有意な差は認めなかった。

各 LR モデルの肝細胞癌予測式を HCC 群と非 HCC 群の各遡り期間で欠測値がない場合の症例データに適用し、各遡り期間における HCC の予測確率 (p 値) を推定した。それぞれの LR モデルにおける各 AUC と AUC の比較結果を表 3 に示す。LRM3 および LRM4 では HCC 群で AUC が -12M から 0M に向け傾向検定で有意な上昇を認めた。また、LRM4 の AUC は、0M と -3M の間に有意差を認めなかったが、0M の AUC は -6M、-9M、-12M の遡り期間と比べ、また、-3M は -9M、-12M の遡り期間と比べ有意に高い AUC であった。一方、LRM3 では、0M の AUC のみ -9M、-12M の遡り期間と比べ有意に高い AUC であり、LRM4 の方が遡り期間での AUC の変化がより明らかと考えられた。LRM4 の式を下に示す。

$$\text{LRM4} = 1 / (1 + \exp(-19.514 + 22.084 \cdot \text{AFP}^{-0.3} - 1.698 \cdot [\text{AFP}/(\text{AST} \times \text{ALT})]^{-0.3} + 4.238 \cdot \ln(\text{AST})))$$

4) 特異度 80%における陽性率の経時的変化

AFP 単独、PIVKA-II 単独、および LRM4 のそれぞれの特異度 80%時のカットオフ値における -12M、-9M、-6M、-3M、0M の経時的な陽性率の比較結果を図 5 に示す。

比較対象とした AFP 単独 (カットオフ値 : 13.5 $\mu\text{g/L}$) での、-12M、-9M、-6M、-3M、診断時の陽性率は、48%、47%、50%、52%、57%であり、PIVKA-II 単独 (カットオフ値 : 40.2 mAU/mL) での陽性率は、15%、13%、20%、43%、38%であった。一方、LRM3 (カットオフ値 : 0.56) では、52%、55%、63%、61%、66%、LRM4 (カットオフ値 : 0.57) では、54%、57%、65%、70%、72%であった。以上の結果から、LRM4 のみ 1 年前から診断時にかけて漸増的に陽性率が上昇し、1 年前から 6 ヶ月前にかけ 11%、6 ヶ月前から診断時にかけ 7%と 1 年間で 18%の経時的な陽性率の上昇を認めた。

4 つのモデルすべての傾向検定で HCC 群の経時的な陽性率の変化は有意であったが、非 HCC 群ではどのモデルにおいても有意ではなかった。

早期ステージの HCC 症例に限定した場合の陽性率を図 5 B に示す。LRM4 による陽性率の上昇は全症例の結果と同様に、1 年前から診断時にかけ漸増的に陽性率が上昇し、また、その間で 20%の陽性率の経時の上昇を認め、最適な診断モデルであると考えられた。

表 2. 各ロジスティック回帰モデルの肝細胞癌の診断性能とモデルの信頼性の比較

model	selected variables	variables (n)	HCC (n)	non-HCC (n)	AIC	training dataset AUC (95% CI)	HCC (n)	non-HCC (n)	validation dataset AUC (95% CI)	P-value
LRM1	*** AFP, PIVKA- II	2	33	40	75.5	0.787 (0.673–0.901)	51	21	0.774 (0.667–0.881)	0.864
LRM2	*** AFP, PIVKA- II, AFP/(AST×ALT)	3	42	28	76.9	0.788 (0.672–0.904)	42	33	0.793 (0.690–0.897)	0.949
LRM3	*** AFP, * AST, ** ALT	3	52	55	110.8	0.845 (0.773–0.917)	40	47	0.852 (0.773–0.930)	0.903
LRM4	*** AFP, *** AFP/(AST×ALT), ** AST	3	47	55	103.5	0.869 (0.794–0.942)	45	47	0.853 (0.781–0.925)	0.768
LRM5	*** AFP/(AST×ALT), AST, *** ALT	3	45	47	111.2	0.830 (0.752–0.908)	47	55	0.833 (0.750–0.916)	0.958

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$

ロジスティック回帰モデルは、0M のトレニングデータセットを使用して、説明変数のさまざまな組み合わせで構築した。HCC 診断の予測における各モデルの診断性能は AUC で示す。また、各モデルを検証データセットに適用し、AUC を計算してその信頼性を評価した。トレーニングデータセットと検証データセット間の AUC の有意差検定の結果を、P 値として右端の列に示す。

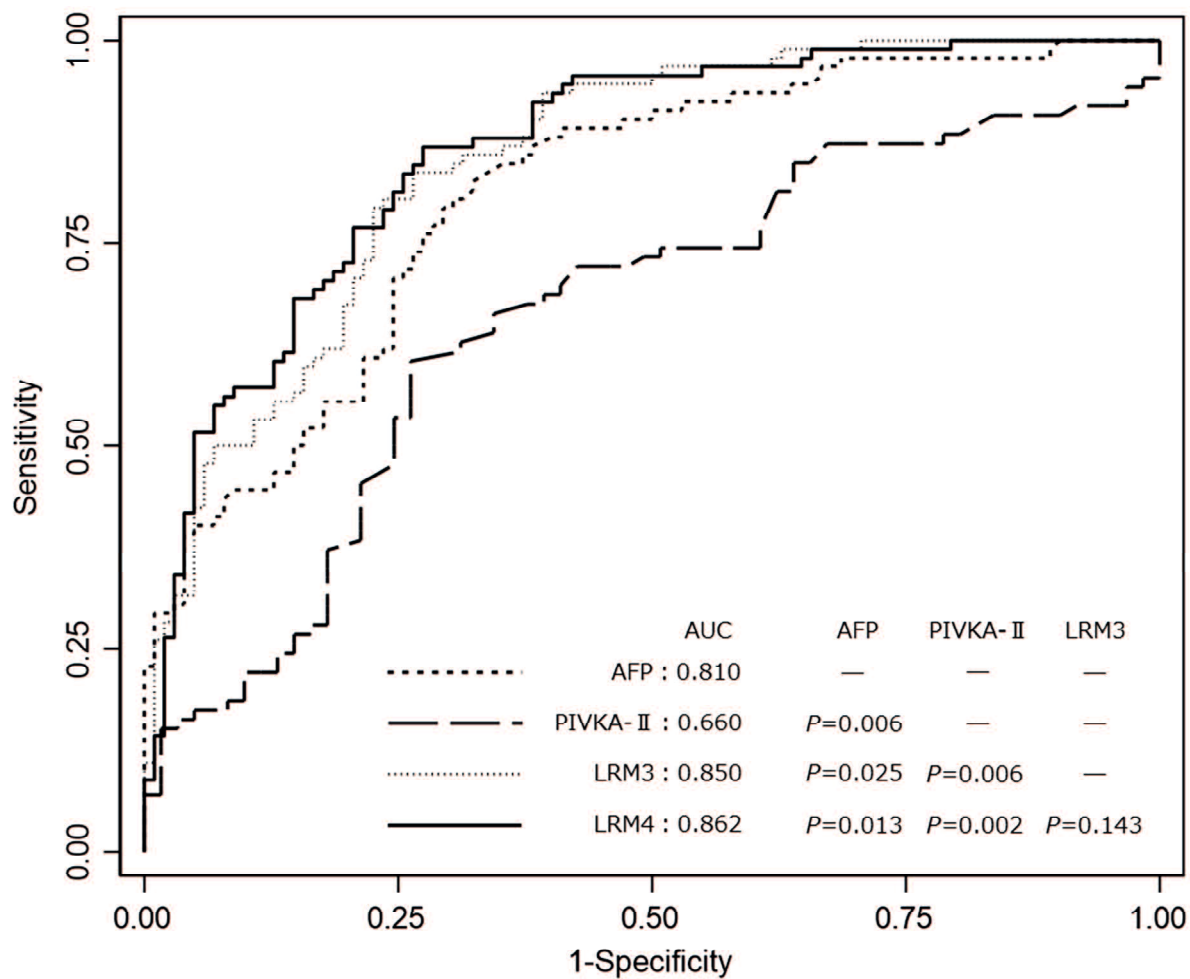


図 4. AFP、PIVKA-II、LRM3、LRM4 の ROC 曲線および AUC の比較

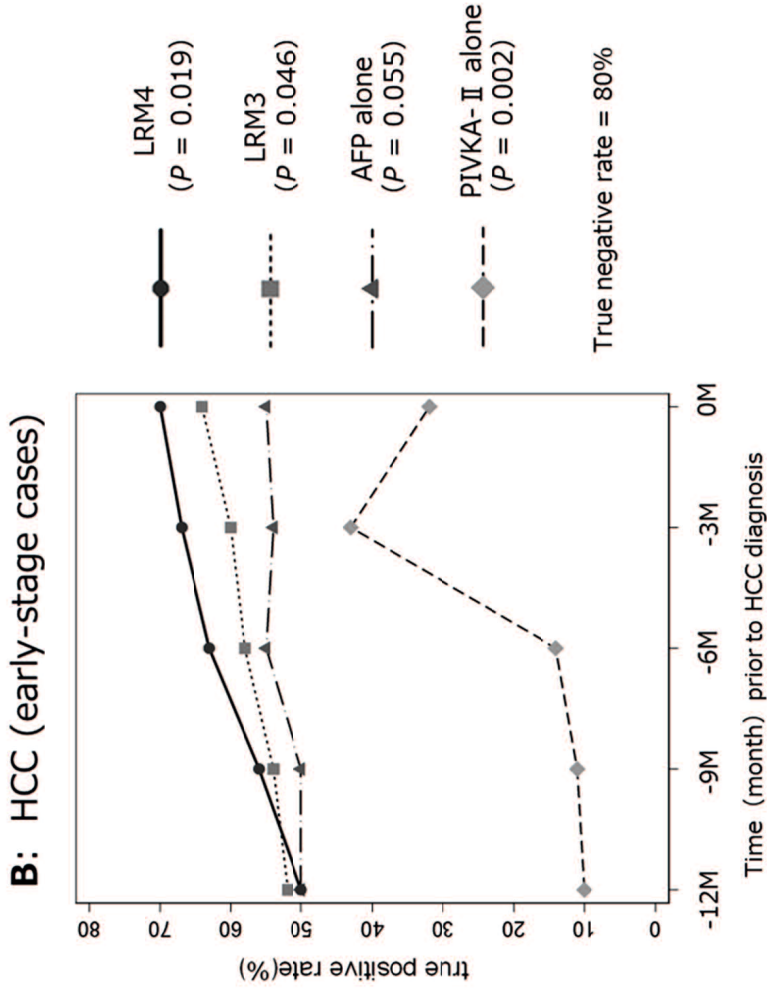
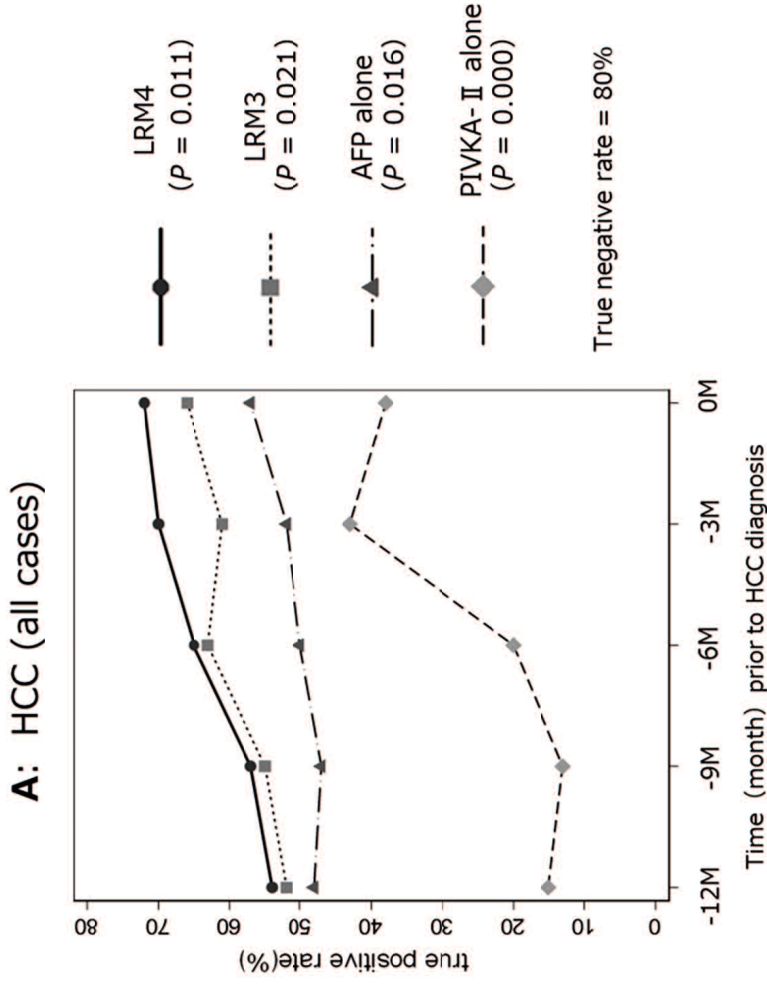
HCC 群と非 HCC 群を区別する診断性能を、AFP 単独、PIVKA-II 単独、LRM3、LRM4 の ROC 曲線によって比較した。HCC の診断能を AUC で示し、各モデル間の AUC の差異の検定結果を P 値で示す。

表 3. 各 LRM の 3 か月 ごと の 遡り 期間 の AUC の 比較

	LRM1	LRM2	LRM3	LRM4	LRM5
AUC	0M 0.773 (0.611 – 0.936)	0.769 (0.606 – 0.932)	0.825 (0.715 – 0.934)	0.849 (0.748 – 0.950)	0.796 (0.677 – 0.915)
(95% CI)	–3M 0.773 (0.609 – 0.938)	0.773 (0.611 – 0.936)	0.792 (0.673 – 0.911)	0.834 (0.726 – 0.941)	0.724 (0.587 – 0.861)
	–6M 0.721 (0.542 – 0.899)	0.737 (0.561 – 0.912)	0.769 (0.642 – 0.896)	0.786 (0.662 – 0.909)	0.750 (0.621 – 0.879)
	–9M 0.794 (0.635 – 0.952)	0.733 (0.555 – 0.911)	0.734 (0.602 – 0.866)	0.752 (0.624 – 0.880)	0.701 (0.559 – 0.843)
	–12M 0.721 (0.535 – 0.906)	0.692 (0.502 – 0.883)	0.684 (0.541 – 0.827)	0.654 (0.506 – 0.802)	0.676 (0.532 – 0.820)
0M vs. –3M	1.000	0.913	0.276	0.502	0.089
0M vs. –6M	0.216	0.444	0.090	*0.044	0.273
0M vs. –9M	0.654	0.334	*0.033	*0.011	0.113
0M vs. –12M	0.403	0.251	*0.034	**0.006	0.063
–3M vs. –6M	0.200	0.244	0.373	0.066	0.467
–3M vs. –9M	0.604	0.243	0.123	*0.024	0.539
–3M vs. –12M	0.399	0.220	0.067	*0.011	0.299
–6M vs. –9M	0.060	0.907	0.321	0.263	0.176
–6M vs. –12M	1.000	0.528	0.131	0.077	0.115
–9M vs. –12M	0.073	0.576	0.327	0.159	0.479

* $P < 0.05$ ** $P < 0.01$

図 5. HCC 群と非 HCC 群の 4 つの予測モデルを使用した経時的な真陽性率の変化



6. 考察

C型慢性肝炎・肝硬変患者でのHCC発症の高リスク群に対する定期的なサーベイランスによる早期診断における腫瘍マーカーの有用性は限定的であり、各種ガイドラインではHCVによるCH/LCでは6ヶ月間隔での超音波等の画像検査との併用が推奨されている²⁴⁻²⁵⁾。超音波による検査精度が施行医の技能に依存し、早期ステージのHCCを検出する感度も限られているため²⁶⁾、これを補完する目的で腫瘍マーカーが活用されている状況にある。一方、HCCの腫瘍容積倍化時間の中央値は3~4ヶ月であることから²⁷⁾、診断時よりもかなり以前から肝細胞癌は存在すると考えられる。しかし、HCCのサーベイランス目的で、腫瘍マーカーを含む一般的な血液検査の経時的な変化から、HCCの早期診断における有用性について検討した報告は少ない。

今回、HCC群における経時的な検査結果の変動を観察したところ、AFP、PIVKA-IIに加え、AFP/(AST×ALT)でも非肝細胞癌群には認めない診断時に向けての有意な経時の上昇変化を認めた。この経時的な上昇は、肝線維化の指標であるFIB-4を含めたマッチング症例で、HCC群のみで見られたことから、肝線維化の影響よりもHCCの影響による上昇であることが示唆された。

今回、これらの腫瘍マーカーに加え、経時的な上昇を認めた一般的な血液検査を加えて多変量ロジスティック回帰による診断モデルを探索的に検討したところ、AFP、AFP/(AST×ALT)、ASTを説明変数としたLRM4で、有意に診断能が改善されることが示された。この診断モデルを過去1年間の3ヶ月毎の検査値に当てはめた結果、陽性率が診断の1年前から診断時に向けて漸増的な上昇を認め、肝細胞癌の影響をより良く捉えたモデルであると考えられた。

さらに1年前からの陽性率の変化を見ると、その陽性率は経時的に上昇し、1年前から6ヶ月前で11%、6ヶ月前から診断時で7%の陽性率の上昇を認めた。

HCC群のSVR(+)の割合が少ないという問題点については、LRM4において説明変数にSVRを追加することによりモデルへの影響を評価したが、SVRは変数としては有意性を認めず、また、診断性能にも影響はなかった。これによりSVRの有無に関わらず、今回のモデルは適用可能であると考えられた。

また、早期ステージに限定したHCCにおいても、LRM4では、陽性率は1年前から診断時で20%の陽性率の上昇を認め、診断性能は全症例の場合と同様であった。したがって、C型慢性肝炎・肝硬変のサーベイランスにおける早期ステージのHCC症例に対しても、LRM4によるサーベイランスは有効であることが示唆された。

一方で、LRM4の他施設における活用性については、AST、ALTの測定法の標準化とAFPの測定法のハーモナイゼーションに依存するが、ASTとALTの測定法のトレーサビリティは認定標準物質を使用して確認でき、また、近年の調査ではAFPの測定法はハーモナイゼーションされてい

ると報告されている²⁸⁾。したがって、LRM4は他の多くの施設においてもHCCのサーベイランスにも適用できると考えられる。

本研究の限界として、第一にHCC群での腫瘍径・個数などの腫瘍の性状以外の状況（肝外転移の有無、腫瘍塞栓症など）を考慮しておらず、また、背景肝の線維化を組織学的ではなくFIB-4で確認していること、第二に後ろ向きに既存の診療情報を活用した研究のため、全ての検査項目が3ヶ月毎の実施がされておらず陽性率の推定に影響した可能性があること、最後に、今回の研究では、比較的少ない症例数で解析しており、トレーニング群と検証群に十分な症例数の割り当てることができなかったが、今回のLRMでは、3つの説明変数しか含まれていないため、LRMの再現性は確保できたものと考えられる。

7. 結語

HCCのサーベイランスにおける臨床検査の診断的有用性について、C型慢性肝炎・肝硬変の患者の経時的な変化に注目した症例対照研究によって評価した。多変量ロジスティック回帰分析(LR)により、AFP、AFP/(AST×ALT)、ASTを説明変数としたLRM4が、HCCと非HCCを鑑別する最適なモデルであった。LRM4による診断時の肝細胞癌予測値を特異度が80%となる値をカットオフ値とし、診断時より3ヶ月単位で1年前まで遡って予測した陽性率は、診断時に向かって漸増的な上昇を示した。この変化は、早期ステージのHCC症例でも同様であった。したがって、LRM4を加えたHCCの定期的なサーベイランスは、その結果が陽性となった患者において、HCCの可能性を疑い、より詳細な画像検査を早期に行う必要性を判断するのに有用であると考えられた。

8. 謝辞

博士課程でご指導をいただいた山口大学大学院医学系研究科医療情報判断学講座の石田 博教授、および本研究でご指導をいただいた臨床検査・腫瘍学講座の山崎 隆弘教授、消化器内科学講座の日高 勲先生、データ解析でご指導をいただいた保健学系学域病態検査学分野の市原 清志名誉教授、データ取得で支援いただいた山口大学医学部附属病院医療情報部の河村 千乃さんに深謝申し上げます。

9. 参考文献

- 1) National Cancer Center Cancer Information Service Vital statistics cancer death data https://ganjoho.jp/reg_stat/statistics/stat/summary.html (2018, accessed 6 November 2020).
- 2) Tateishi R, Uchino K, Fujiwara N, et al. A nationwide survey on non-B, non-C hepatocellular carcinoma in Japan: 2011-2015 update. *J Gastroenterol* 2018; 54: 367-376.
- 3) Kanwal F, Kramer J, Asch SM, et al. Risk of hepatocellular cancer in HCV patients treated with direct-acting antiviral agents. *Gastroenterology* 2017; 153: 996-1005.
- 4) Toyoda H, Tada T, Takaguchi K, et al. Differences in background characteristics of patients with chronic hepatitis C who achieved sustained virologic response with interferon-free versus interferon-based therapy and the risk of developing hepatocellular carcinoma after eradication of hepatitis C virus in Japan. *J Viral Hepat* 2017; 24: 472-476.
- 5) Norio H, Treatment method for hepatitis C, supervised by the Japanese Society of Liver Disease, Chronic Hepatitis Medical Treatment Manual, p60-68 Igaku-Shoin, Tokyo 2001; 5.
- 6) Yamashita N, Ohho A, Yamasaki A, et al. Hepatocarcinogenesis in chronic hepatitis C patients achieving a sustained virological response to interferon: significance of lifelong periodic cancer screening for improving outcomes. *J Gastroenterol* 2014; 49: 1504-1513.
- 7) Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2004; 130: 417-422.
- 8) Chen JG, Parkin DM, Chen QC, et al. Screening for liver cancer: results of a randomized controlled trial in Qidong, China. *J Med Screen* 2003; 10: 204-209.
- 9) Arguedas MR, Chen VK, Eloubeidi MA, et al. Screening for hepatocellular carcinoma in patients with hepatitis C cirrhosis: a cost-utility analysis. *Am J Gastroenterol* 2003; 98: 679-690.
- 10) Bolondi LS, Sofia S, Siringo S, et al. Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma: a cost effectiveness analysis. *Gut* 2001; 48: 251-259.
- 11) Coon JT, Rogers G, Hewson P, et al. Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. *Br J Cancer* 2008; 98: 1166-1175.
- 12) Giorgio A, Calisti G, di Sarno A, Farella N, de Stefano G, Scognamiglio U, Giorgio V. Characterization of dysplastic nodules, early hepatocellular carcinoma and progressed hepatocellular carcinoma in cirrhosis with contrast-enhanced ultrasound. *Anticancer Res* 2011; 11: 3977-3982.
- 13) Wu W, Chen M, Yan K, et al. Evaluation of contrast-enhanced ultrasound for diagnosis of dysplastic nodules with a focus of hepatocellular carcinoma in liver cirrhosis patients. *Chin J*

Cancer Res 2015;1: 83-89.

- 14) Di Martino M, Marin D, Guerrisi A, et al. Intraindividual comparison of gadoxetate disodium-enhanced MR imaging and 64-section multidetector CT in the detection of hepatocellular carcinoma in patients with cirrhosis. *Radiology* 2010;3: 806-816.
- 15) Trinchet JC, Chaffaut C, Bourcier V, et al. Ultrasonographic surveillance of hepatocellular carcinoma in cirrhosis: a randomized trial comparing 3-and 6-month periodicities. *Hepatology* 2011; 54: 1987-1997.
- 16) Wang JH, Chang KC, Kee KM, et al. Hepatocellular carcinoma surveillance at 4- vs. 12-month intervals for patients with chronic viral hepatitis: a randomized study in community. *Am J Gastroenterol* 2013; 108: 416-424.
- 17) Li C, Zhang Z, Zhang P, et al. Diagnostic accuracy of des-gamma-carboxy prothrombin versus alpha-fetoprotein for hepatocellular carcinoma: a systematic review. *Hepatol* 2014; 44: E11-E25.
- 18) Lok AS, Sterling RK, Everhart JE, et al. Des-gamma-carboxy prothrombin and alpha-fetoprotein as biomarkers for the early detection of hepatocellular carcinoma. *Gastroenterology* 2010; 138: 493-502.
- 19) Choi J, Kim GA, Han S, et al. Longitudinal assessment of three serum biomarkers to detect very early-stage hepatocellular carcinoma. *Hepatology* 2019; 69: 1855-1857.
- 20) Liu X, Meng J, Xu H, et al. Alpha-fetoprotein to transaminase ratio to higher diagnostic efficacy for hepatocellular carcinoma. *Medicine* 2019; 98.
- 21) Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006; 43: 1317-1325.
- 22) Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat* 1985; 39: 33-38.
- 23) Shintani A. Multivariate analysis-how to select explanatory variables. *Medical statistics that can be used from today*. Tokyo: Igaku-Shoin; 2015: 101-113.
- 24) Heimbach JK, Kulik LM, Finn RS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* 2018; 67: 358-380.
- 25) European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu; European Association for the Study of the Liver. *EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma*. *J Hepatol* 2018 Jul; 69: 182-236.
- 26) Singal A, Volk ML, Waljee A, et al. Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. *Aliment Pharmacol Ther* 2009; 30: 37-47.

- 27) An C, Choi YA, Choi D, et al. Growth rate of early-stage hepatocellular carcinoma in patients with chronic liver disease. *Clin Mol Hepatol* 2015; 21: 279-286.
- 28) Oikawa S, Horiuchi Y, Ikeda M, Hishinuma A. Recent progress of AFP and PIVKA-II assays. *J Lab Clin Med* 2015; 59: 602-609.