Kathryn J. Fowler, MD
An Tang, MD, MSc, FRCPC
Cynthia Santillan, MD
Mythreyi Bhargavan-Chatfield, PhD
Jay Heiken, MD
Reena C. Jha, MD
Jeffrey Weinreb, MD
Hero Hussain, MD
Donald G. Mitchell, MD
Mustafa R. Bashir, MD
Eduardo A. C. Costa, MD
Guilherme M. Cunha, MD
Laura Coombs, PhD
Tanya Wolfson, MA
Anthony C. Gamst, PhD
Giuseppe Brancatelli, MD
Benjamin Yeh, MD
Claude B. Sirlin, MD

ORIGINAL RESEARCH ■ GASTROINTESTINAL

Interreader Reliability of LI-RADS Version 2014 Algorithm and Imaging Features for Diagnosis of Hepatocellular Carcinoma: A Large International Multireader Study[1]

Radiology: Volume 286: Number 1—January 2018 ▪ *radiology.rsna.org*

Reunião Bibliográfica

Nuno Pereira da Silva - 17 Janeiro de 2018

Diretor de Serviço: Prof. Doutor Filipe Caseiro Alves

Objetivos

- Determinar, num estudo multicêntrico alargado, a *interreader reliablity* (IRR) das categorias do LI-RADS v2014 (*Liver Imaging Reporting and Data System*, versão 2014) e das características imagiológicas major observadas em TC e RM.

- Determinar o potencial efeito de variáveis demográficas dos radiologistas na concordância da interpretação de um conjunto de imagens pré-selecionadas.

- Objetivo secundário – avaliar o contributo das características auxiliares e regras de "desempate" na categorização.

Introdução

Achados estabelecidos no CHC:

- Marcado aumento do realce na fase arterial;
- “Washout”;
- Realce da cápsula;
- Tamanho e aumento dimensional.

Inúmeros sistemas de diagnóstico de CHC: AASLD, OPTN, APASL, EASL…

➢ **LI-RADS**

- organização em categorias, léxico normalizado e características major e auxiliares ao diagnóstico.
- promove a standardização e reprodutibilidade entre instituições e radiologistas.

Introdução

Table 1

Diagnostic Imaging Criteria for Definite HCC

Diagnostic System	Imaging Criteria for Definite HCC	Specific Features
AASLD*	Lesion seen at antecedent US. At CT or MR imaging, lesion diameter >1 cm indicates APHE and washout	Applied in context of US surveillance.
OPTN†	At CT or MR imaging, for lesions ≥1 cm and <2 cm, one of the following: *(a)* APHE, washout, and capsule growth or *(b)* APHE and ≥50% growth in 6 months or less. For lesions ≥2 cm and ≤5 cm, one of the following: *(a)* APHE and washout, *(b)* APHE and capsule, or *(c)* APHE and threshold growth	Applied in context of MELD, exception point allocation for transplant candidates.
LI-RADS‡	For lesions 1.0 to 1.9 cm, APHE and two of the following: *(a)* washout, capsule growth, or threshold growth; *(b)* lesion seen at antecedent US and APHE and washout; or *(c)* APHE and ≥50% growth in 6 months or less. For lesions ≥2 cm, APHE and at least one of the following: washout, capsule growth, or threshold growth	Applied in all patients at risk. Additional considerations include ancillary features and tie-breaking rules§ for final categorization (LR-1 to LR-5, LR-5V, LR-M)

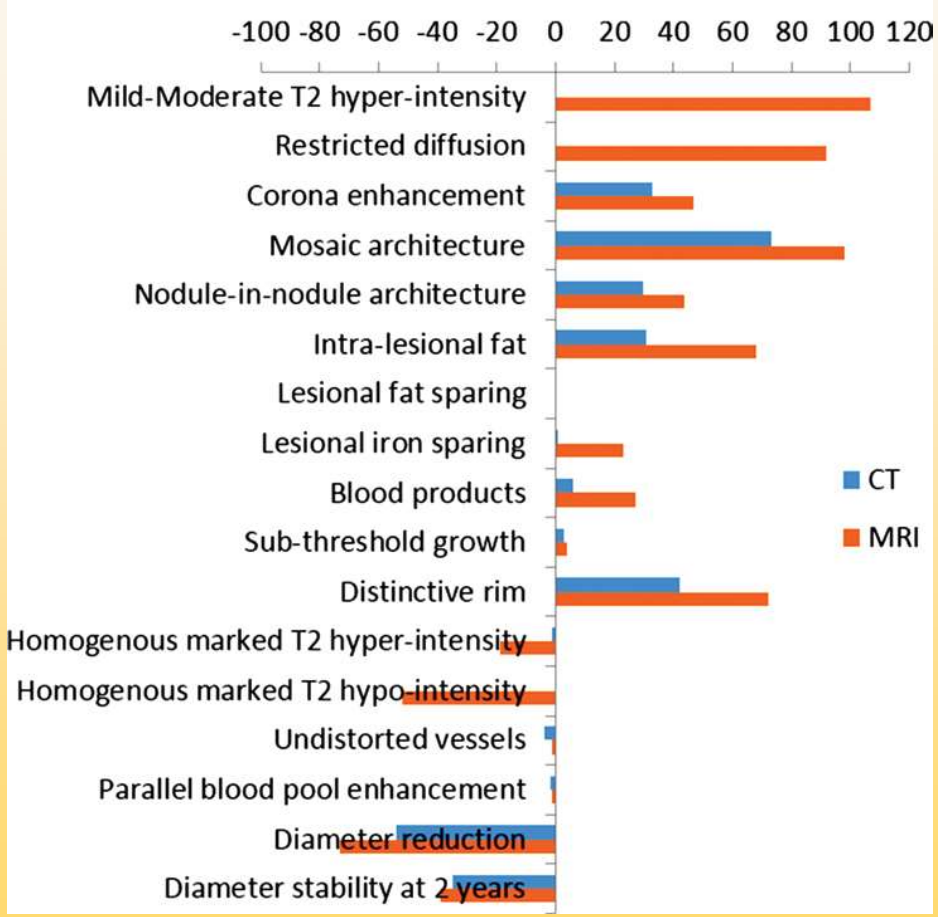
LI-RADS (Version 2014) - Categorias

Table 2

LI-RADS Version 2014 Categories and Definitions

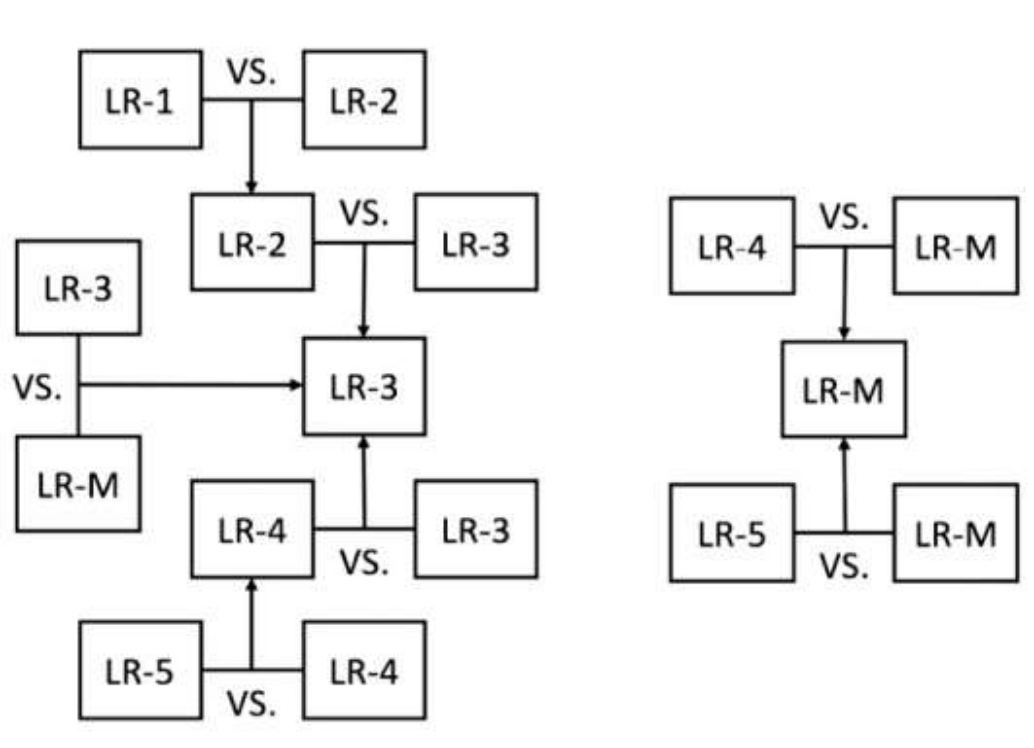
LI-RADS Category	Definition
LR-1	Definitely benign
LR-2	Probably benign
LR-3	Intermediate probability for HCC
LR-4	Probably HCC
LR-5	Definitely HCC
LR-5V	Definite tumor in vein
LR-M	Probably malignant, not specific for HCC

LI-RADS (Version 2014) – Características Auxiliares



- Não permitem a categorização das lesões isoladamente;
- Podem ser utilizadas para fazer o *upgrade* de lesões ou *downgrade* de lesões;
- *Upgrade* não pode exceder LR-4 (para preservar a especificidade diagnóstica do CHC).

LI-RADS (Version 2014) – Regras de “Desempate”



- Utilizadas para ajustar a categoria quando subsiste a incerteza mesmo após aplicação do algoritmo ou das “características auxiliares”.

Introdução

Table 6

Summary of Agreement from Literature

Study and LI-RADS Version	No. of Readers, No. of Lesions, and Case Selection Criteria	Modality and Format of Image Sets
Davenport M (LI-RADS version 2013.1)†	10 readers (5 expert, 5 novice), 100 lesions, consecutive cases scored by research team from 2011–2013 to meet quota of 20 per LI-RADS category	MR imaging ECA phases, scrollable image sets
Barth et al (LI-RADS version 2013.1)‡	Four readers (8–12 years of experience), 104 lesions, consecutive cases from database of LI-RADS categories screened for quota of about 15 per LI-RADS category	MR imaging with ECA, scrollable image sets, and a screenshot showing observation
Zhang et al (LI-RADS version not specified)§	Two readers (6 years of postfellowship experience), 118 lesions, consecutive cases from existing database selected to include cases with both CT and MR images available	MR imaging with ECA, CT, and scrollable image sets
Ehman et al (LI-RADS version 2014)‖	Two readers (experience not specified), 184 lesions, consecutive cases with histopathology proof of diagnosis of HCC	MR imaging with ECA, CT, scrollable image sets
Sofue et al (LI-RADS version 2013.1)#	Two readers (7 and 6 years of experience), 111 lesions, consecutive cases with hypervascular nodules by report	MR imaging with EC phases, scrollable image sets
Bashir et al (LI-RADS version 2013.1)**	Three readers (2, 4, and 12 years of experience), 200 lesions, consecutive cases with hypervascular nodules by report	MR imaging with ECA, CT, MR imaging with HBA, scrollable image sets
Current study (LI-RADS version 2014)	113 readers (average 11 readers, diverse background), 380 lesions, selected cases from multiple institutions submitted to a liver atlas	MR imaging with ECA, CT, PowerPoint image captures

Materiais e métodos – “Atlas Imagiológico”

- Construído um “**atlas**” com 382 casos correspondendo a observações por TC e RM (por nove radiologistas com 9-35 anos de experiência, membros do Comité LI-RADS):
  - Imagens clássicas e achados complexos
  - Lesões malignas e benignas

- **TC**:

- protocolo sem contraste, com contaste nas fases arterial, portal e tardia

- em imagens dúbias foram ainda fornecidos reconstruções sagitais, coronais e radiodensidade em UH

- **RM**:

- obtidas imagens com protocolo normal de fígado (*T2-weighted imaging, diffusion-weighted imaging, unenhanced T1-weighted in- and opposed- phase imaging, and dynamic unenhanced and contrast-enhanced imaging in the late arterial, portal venous,and delayed phases, along with arterial subtraction imaging*)

- só foram utilizados contrastes extracelulares

Materiais e métodos – “Image Sets”

- Todos os casos (n=382) foram revistos em consenso por dois membros do Comité LI-RADS, sendo-lhes atribuído uma categoria:
  21 LR-1; 56 LR-2; 96 LR-3; 74 LR-4; 91 LR-5; 28 LR-5V; 16 LR-M

- Divididos em 10 conjuntos, cada um com 38-40 casos, selecionados aleatoriamente mas com igual representação de TC e RM, assim como das categorias de LR-1 a LR-5, LR-5V, LR-M.

- Cada radiologista preencheu um inquérito quanto ao **grau de familiaridade com o LI-RADS**, o número de **anos como especialista**, o grau de **experiência com imagiologia hepática** e com **outros algoritmos imagiológicos**.

Materiais e métodos – “Readers”

- Participaram 113 radiologistas;
- Selecionados pela experiência com imagem hepática, por pertencerem a instituições filiadas com o LI-RADS Steering Comitee ou por recomendação por outros radiologistas.
- Largo espectro de radiologistas (peritos e não-peritos em imagem hepática) de múltiplas nacionalidades.

Materiais e métodos – “Readers”

Demographic Information	No. of Readers (n = 113)
Year completed residency	
1970–1989	18 (16)
1990–2004	49 (43)
2005–2014	46 (41)
Body fellowship trained	
Yes	89 (79)
No	23 (21)
Practice type	
Academic	74 (65)
Combined	34 (30)
Private	5 (4)
Liver transplant center	
Yes	75 (66)
No	38 (34)
Geographic region	
Asia	15 (13)
Europe	20 (18)
North America	68 (60)
Oceana	4 (4)
South America	6 (5)

Expertise in liver imaging	
I am a specialist	99 (87)
I read some liver images	13 (12)
I read very few liver images	1 (1)
Systems I currently use in practice*	
AASLD	48 (42)
APASL	3 (3)
EASL	23 (20)
JSH	5 (4)
LI-RADS	40 (35)
OPTN	42 (37)
No formal system	20 (18)
LI-RADS familiarity	
Not familiar at all	9 (8)
Somewhat familiar	65 (58)
Very familiar	39 (35)

Materiais e métodos – Análise estatística

- Utilizado software *R* (versão 3.3.1)
- Realizada estatística descritiva das variáveis demográficas e das características auxiliares e regras de “desempate”
- Análise de regressão foi utilizada para estimar o Coeficiente de Correlação Intraclasse (CCI) enquanto indicador da **concordância inter-observador (IRR):**
    - 0,20 concordância baixa
    - 0,21-0,40 concordância aceitável
    - 0,41-0,60 concordância moderada
    - 0,61-0,80 concordância substancial
    - 0,81-1,00 concordância quase perfeita

Materiais e métodos – Análise estatística

- Foram testadas quatro estratificações do IRR das categorias de LI-RADS:
  (a) LR-1/LR-2, LR-3, LR-4, LR-5/LR-5V/LR-M;
  (b) LR-1/LR-2, LR-3, LR-4, LR-5/LR-5V;
  (c) LR-1/LR-2, LR-3, LR-4/LR-5/LR-5V/LR-M;
  (d) LR-1/LR-2, LR-3, LR-4/LR-5/LR-5V.

LR-1 e LR-2 foram agrupados por representarem lesões benignas que não requerem follow-up.

LR-5 e LR-5V foram agrupados por representarem lesões que são "definitivamente CHC" não requerendo biópsia.

Em duas estratificações LR-4 foi agrupado com LR-5 e LR-5V para formar a categoria de lesões "provavelmente CHC".

LR-M não segue a probabilidade ordinal da outras categorias pelo que o CCI foi calculado com e sem inclusão desta variável.

Materiais e métodos – Análise estatística

- Foi também utilizado um modelo linear generalizado para o cálculo do CCI para o IRR relativo às <u>características major.</u>

- O IRR foi comparado entre características demográficas reportadas pelos radiologistas participantes, tendo quatro sido dicotomizadas e examinadas em análises separadas:
    - “Grau de experiência com imagem hepática”,
    - “Grau de familiaridade com o LI-RADS”,
    - “Número de anos como especialista”,
    - “Tipo de prática”.

Resultados – Concordância entre categorias de LI-RADS

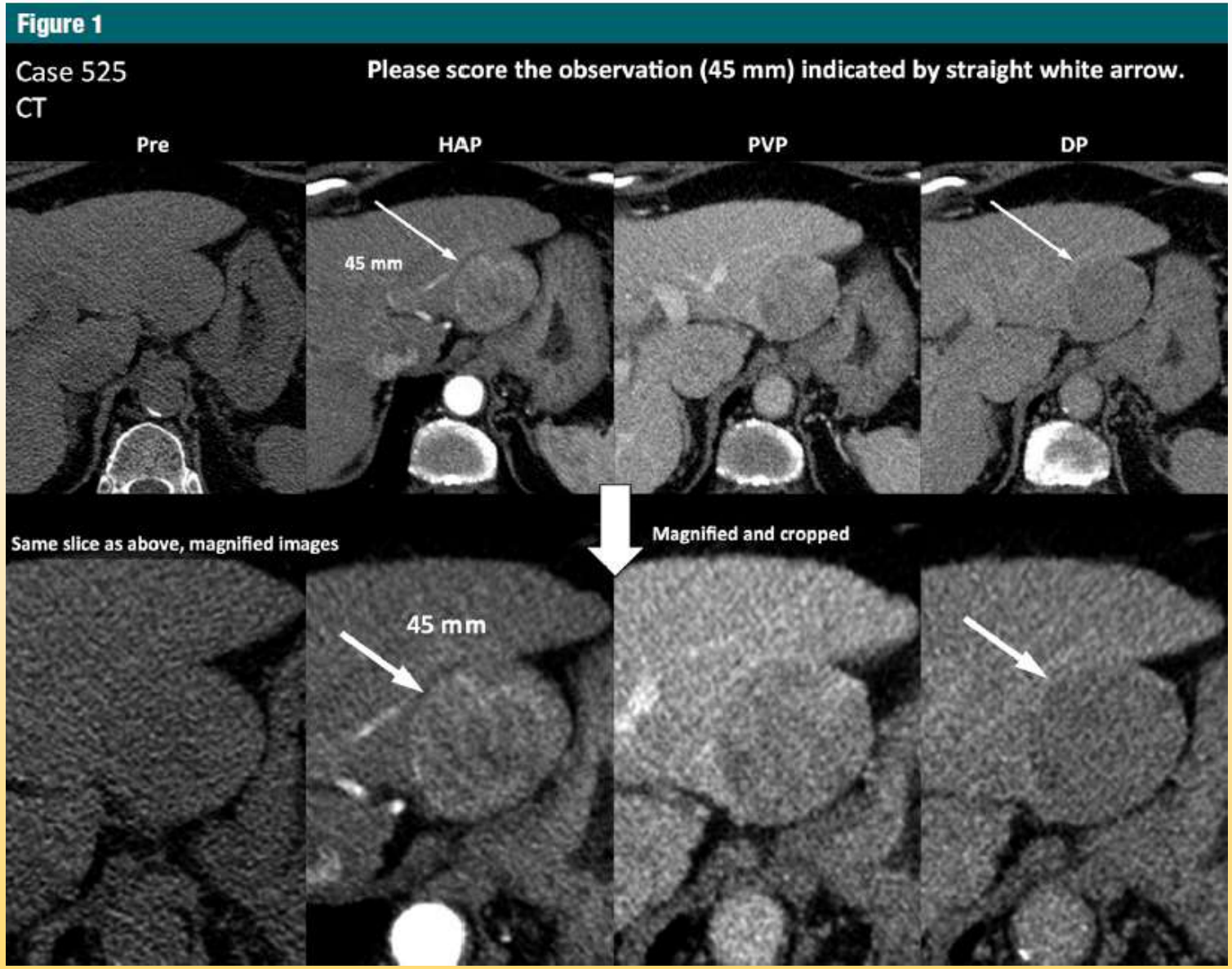
Table 4

Interrater Reliability for Assigning LI-RADS Categories and Major Features

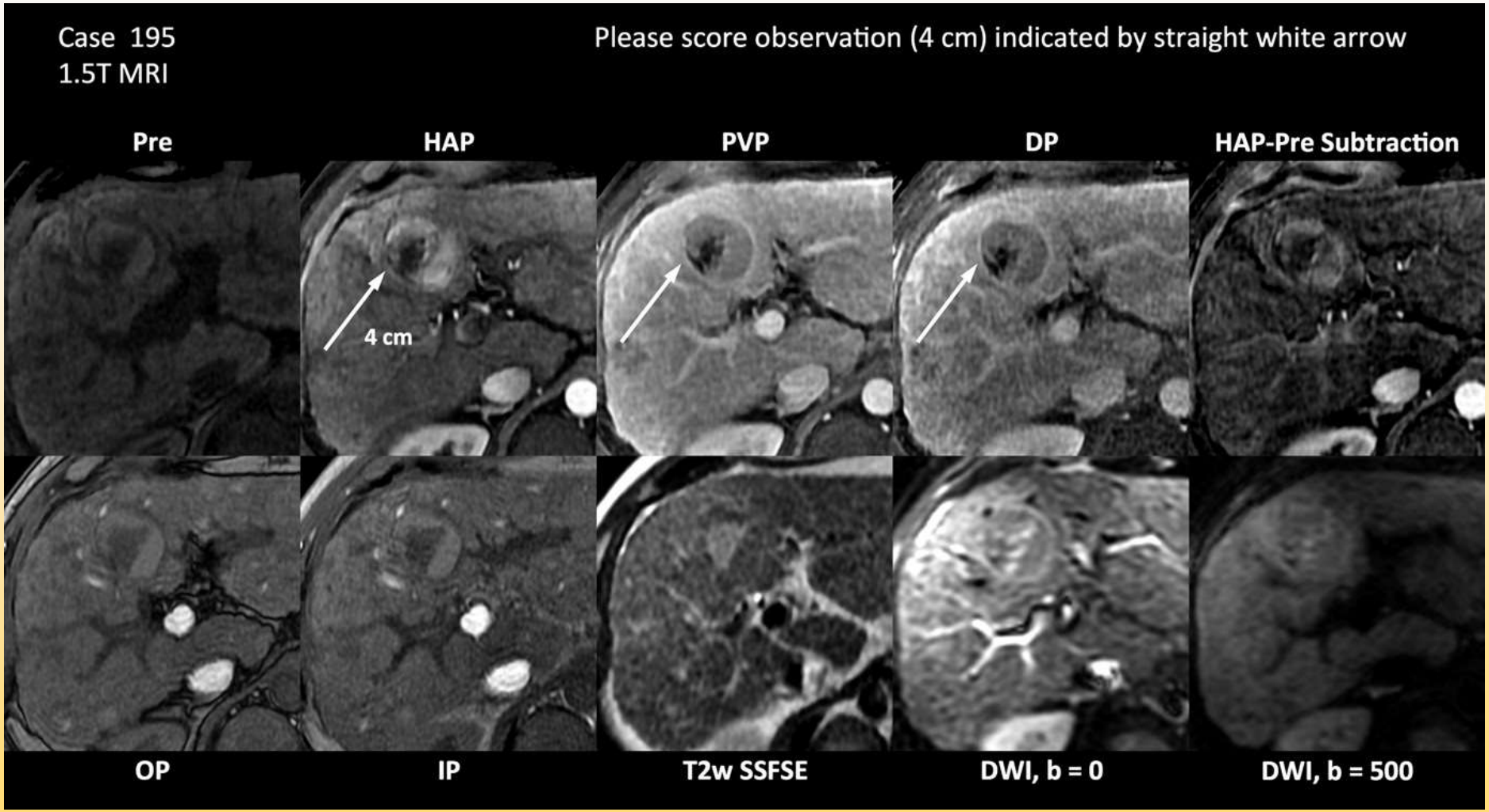
Finding	CT	MR Imaging	Overall
LI-RADS category			
LR-1/LR-2, LR-3, LR-4, LR-5/LR-5V, LR-M	0.67 (0.61, 0.71)	0.73 (0.68, 0.77)	0.70 (0.67, 0.73)
LR-1/LR-2, LR-3, LR-4, LR-5/LR-5V	0.68 (0.62, 0.73)	0.73 (0.68, 0.77)	0.71 (0.67, 0.74)
LR-1/LR-2, LR-3, LR-4/LR-5/LR-5V, LR-M	0.61 (0.56, 0.66)	0.68 (0.63, 0.73)	0.65 (0.62, 0.69)
LR-1/LR-2, LR-3, LR-4/LR-5/LR-5V	0.63 (0.57, 0.68)	0.69 (0.64, 0.73)	0.67 (0.62, 0.70)
Major features			
APHE	0.86 (0.82, 0.91)	0.88 (0.83, 0.94)	0.87 (0.84, 0.90)
Washout appearance	0.84 (0.80, 0.90)	0.84 (0.80, 0.88)	0.85 (0.81, 0.88)
Capsule appearance	0.79 (0.72, 0.85)	0.88 (0.83, 0.92)	0.84 (0.80, 0.87)

Note.—Data are ICCs. Data in parentheses are 95% CIs. LR-M cases were included in major features analyses.

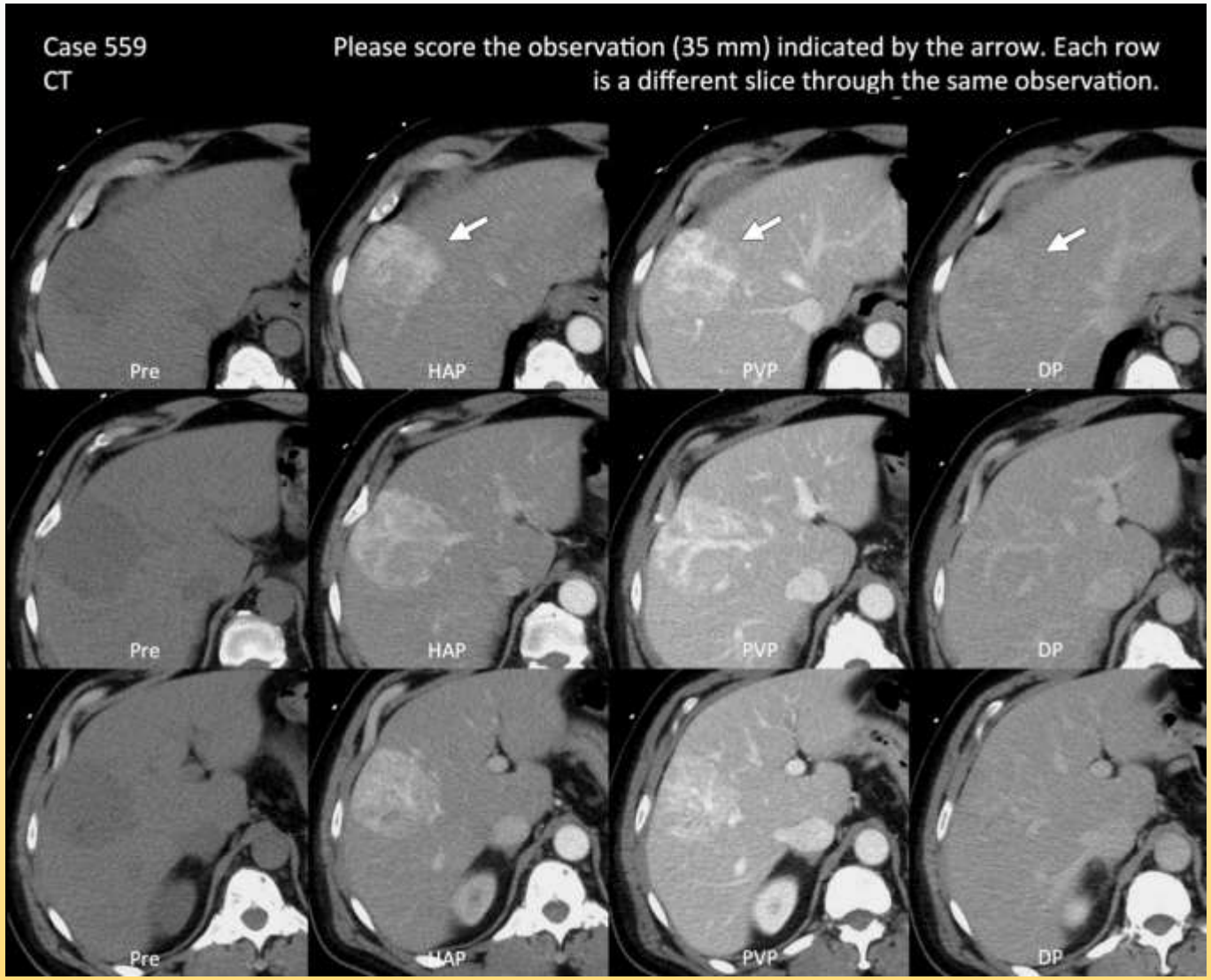
- CCI mais alto para RM;
- CCI mais alto quando LR-4 e LR-5 não estão agrupadas;
- Exclusão de LR-M não teve impacto;



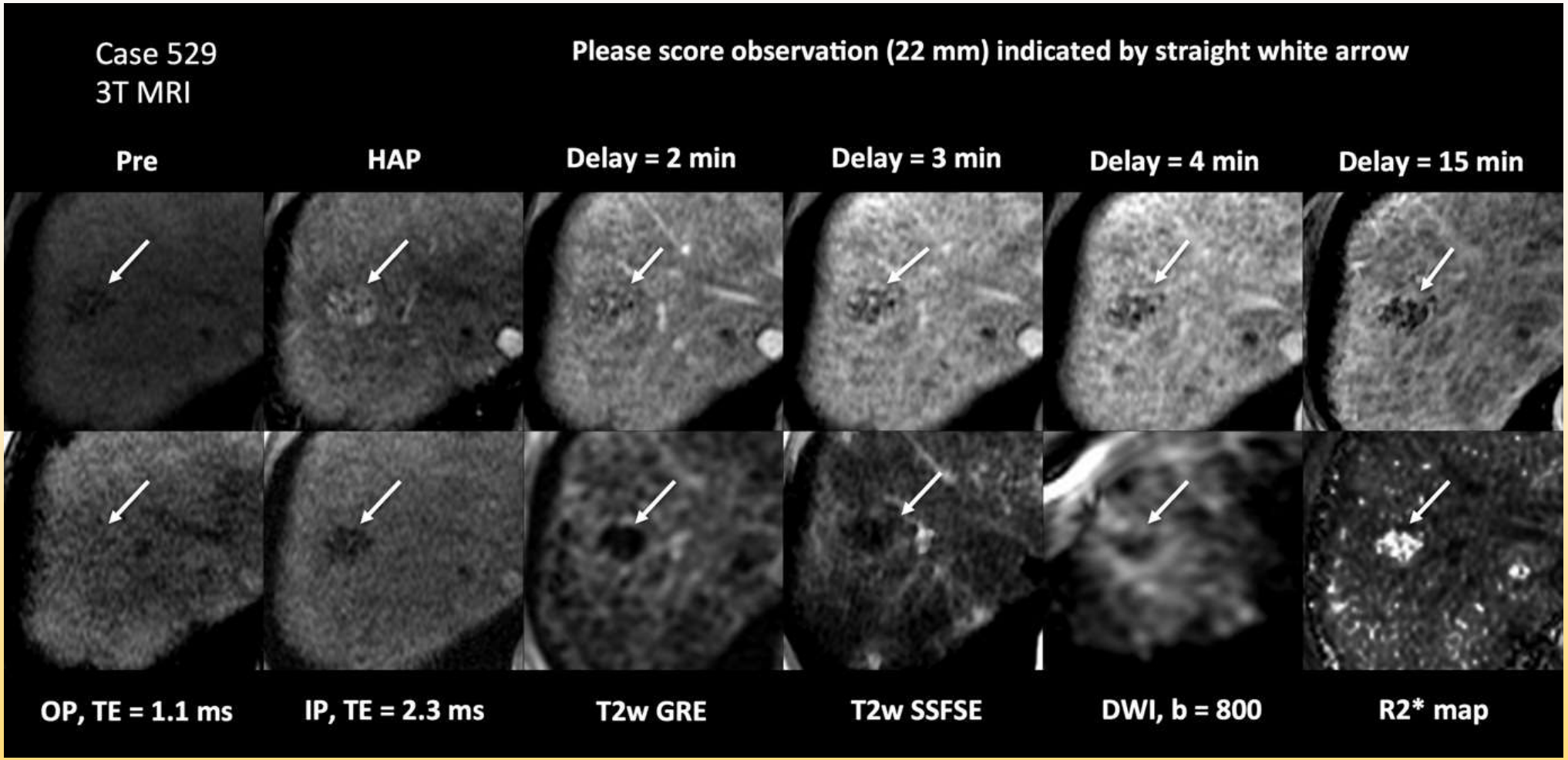
Exemplo de caso TC com **alta concordância** (100% classificou como LR-5)



Exemplo de caso RM com **alta concordância** (100% classificou como LR-5)

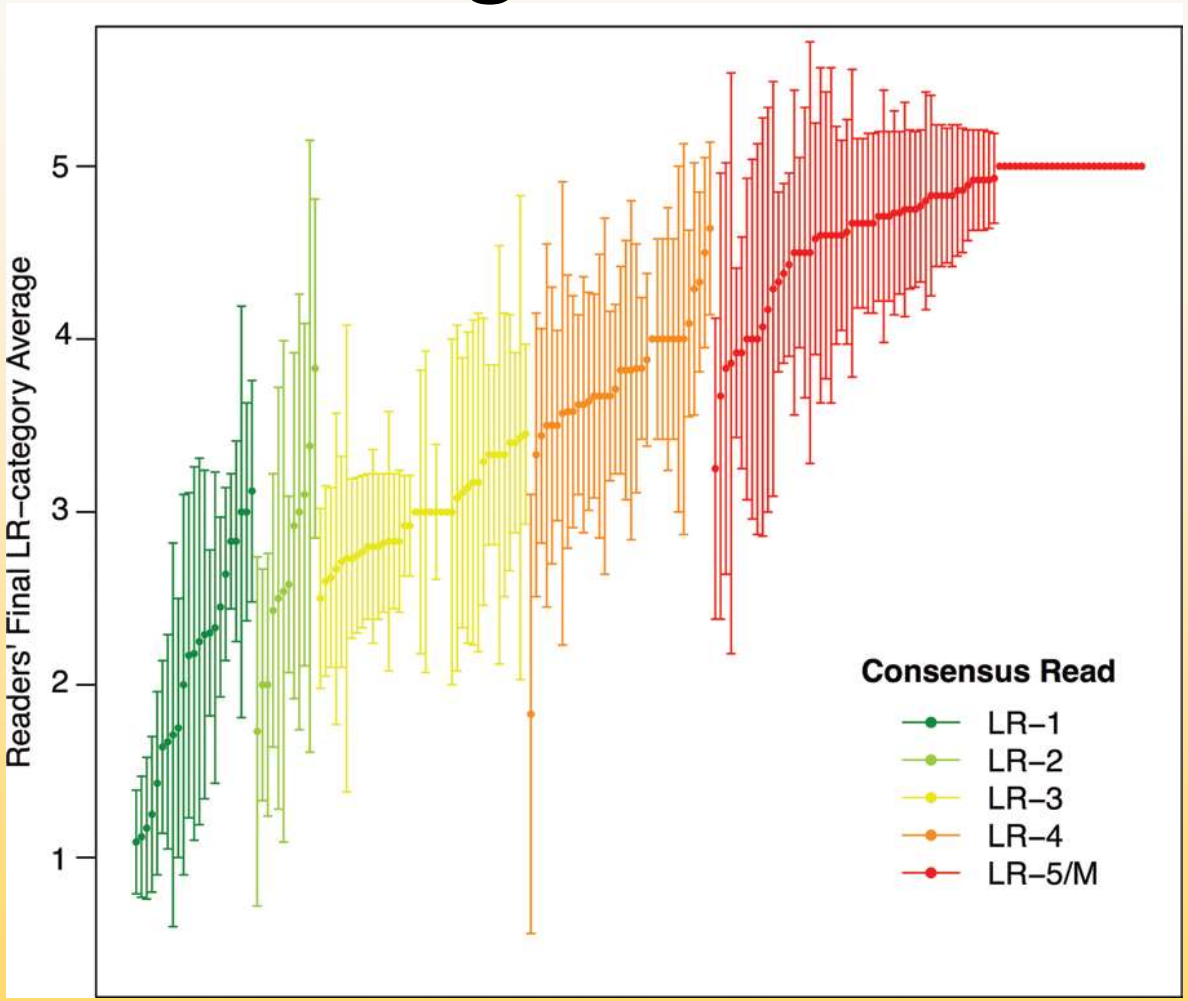


Exemplo de caso TC com **baixa concordância** (Consenso do comité LI-RADS classificou como LR-M; categoria mais frequente LR-5 c/ 33%)

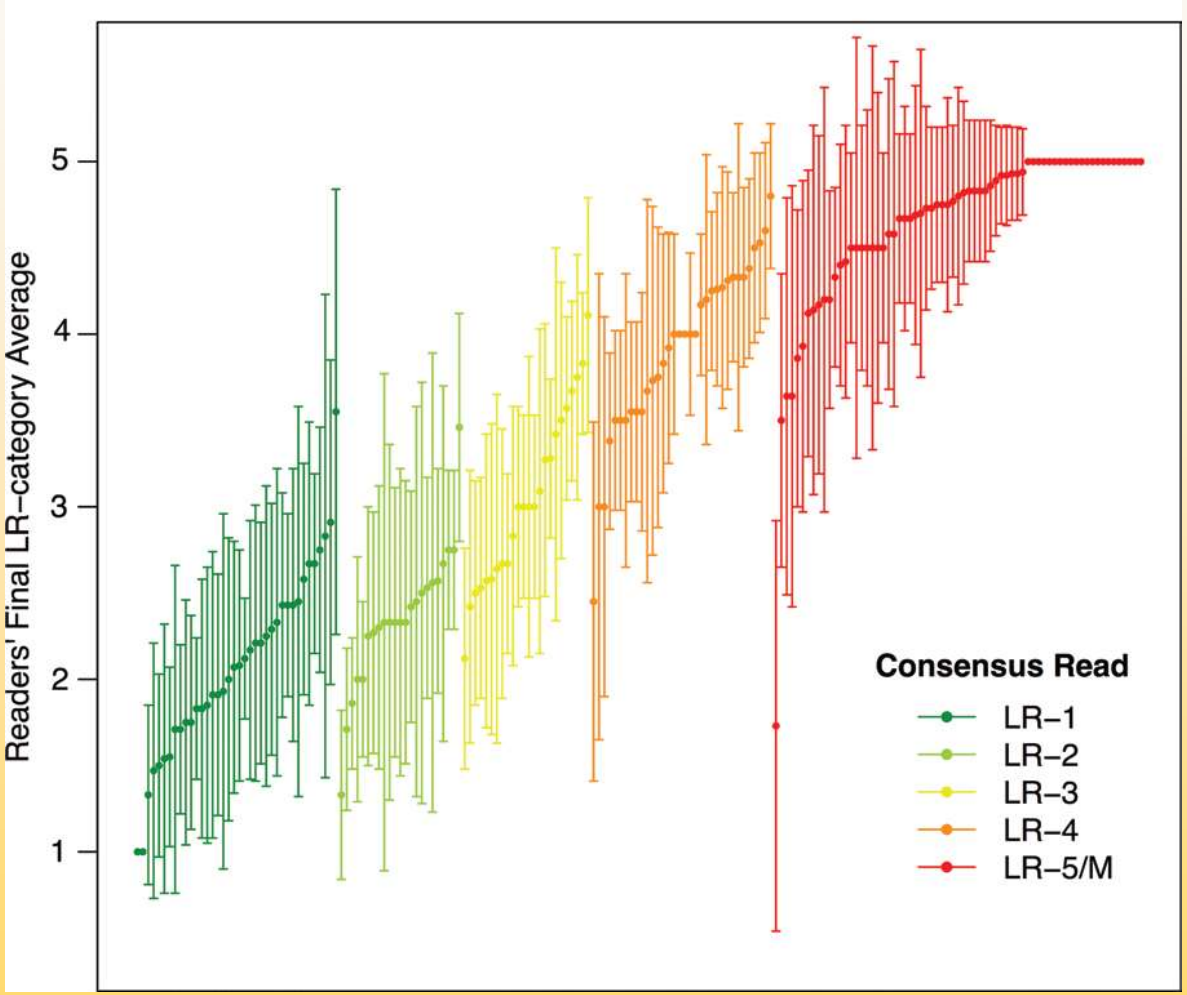


Exemplo de caso RM com **baixa concordância** (Categoria mais frequente foi LR-5, mas apenas para 36% dos leitores)

Distribuição e “scores” médios para conjuntos de imagens



TC



RM

Resultados – Concordância entre Características Major

Table 4

Interrater Reliability for Assigning LI-RADS Categories and Major Features

Finding	CT	MR Imaging	Overall
LI-RADS category			
LR-1/LR-2, LR-3, LR-4, LR-5/LR-5V, LR-M	0.67 (0.61, 0.71)	0.73 (0.68, 0.77)	0.70 (0.67, 0.73)
LR-1/LR-2, LR-3, LR-4, LR-5/LR-5V	0.68 (0.62, 0.73)	0.73 (0.68, 0.77)	0.71 (0.67, 0.74)
LR-1/LR-2, LR-3, LR-4/LR-5/LR-5V, LR-M	0.61 (0.56, 0.66)	0.68 (0.63, 0.73)	0.65 (0.62, 0.69)
LR-1/LR-2, LR-3, LR-4/LR-5/LR-5V	0.63 (0.57, 0.68)	0.69 (0.64, 0.73)	0.67 (0.62, 0.70)
Major features			
APHE	0.86 (0.82, 0.91)	0.88 (0.83, 0.94)	0.87 (0.84, 0.90)
Washout appearance	0.84 (0.80, 0.90)	0.84 (0.80, 0.88)	0.85 (0.81, 0.88)
Capsule appearance	0.79 (0.72, 0.85)	0.88 (0.83, 0.92)	0.84 (0.80, 0.87)

Note.—Data are ICCs. Data in parentheses are 95% CIs. LR-M cases were included in major features analyses.

- CCI elevado para Aumento do Realce na Fase Arterial, “Washout” e para realce da cápsula.
- CCI ligeiramente mais elevado para RM
- Não foi analisado o diâmetro ou aumento dimensional das lesões.

Resultados – Concordância entre Características Demográficas

Table 5

Agreement on LI-RADS Category and Major Features by Reader Demographic Strata

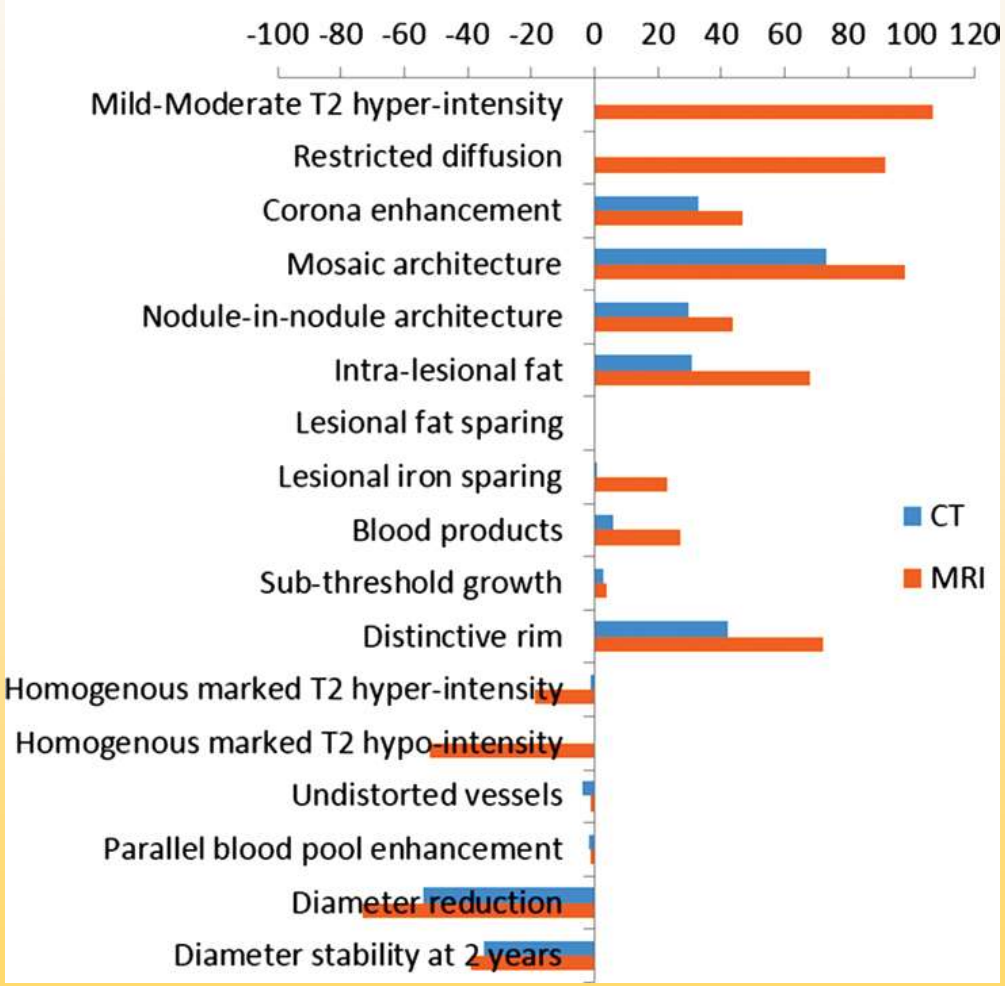
Practice Type	Overall LI-RADS Category	APHE	Capsule	Washout Appearance
Academic	0.71 (0.67, 0.73)	0.86 (0.82, 0.90)	0.84 (0.80, 0.89)	0.85 (0.81, 0.89)
Community or mixed	0.69 (0.66, 0.72)	0.88 (0.82, 0.93)	0.83 (0.77, 0.89)	0.83 (0.77, 0.89)
P value	.050*	.613	.783	.680
Liver expertise				
Yes	0.70 (0.66, 0.73)	0.88 (0.84, 0.91)	0.83 (0.79, 0.87)	0.86 (0.82, 0.89)
Some or none	0.70 (0.66, 0.73)	0.81 (0.69, 0.89)	0.94 (0.74, 1.0)	0.75 (0.63, 0.85)
P value	.826	.134	.248	.0392*
LI-RADS familiarity				
Very familiar	0.70 (0.67, 0.74)	0.88 (0.81, 0.97)	0.83 (0.76, 0.89)	0.92 (0.86, 0.97)
Somewhat or not familiar	0.70 (0.67, 0.73)	0.87 (0.83, 0.90)	0.84 (0.79, 0.89)	0.82 (0.78, 0.86)
P value	.264	.820	.707	.0072*
Years of practice				
≥6 years	0.70 (0.66, 0.73)	0.86 (0.83, 0.90)	0.84 (0.80, 0.88)	0.84 (0.80, 0.87)
<6 years	0.71 (0.67, 0.74)	0.93 (0.83, 0.99)	0.85 (0.72, 0.98)	0.91 (0.81, 0.98)
P value	.852	.198	.911	.204

- Diferenças de CCI ligeiramente mais elevadas nos leitores “não académicos vs académicos”

- CCI não eram significativamente afetados pelo grau de familiaridade com o LI-RADS, o número de anos como especialista, o grau de experiência em imagem hepática (excepto na caracterização do “washout”)

Resultados – Concordância entre Características Auxiliares e Regras de "Desempate"

- Foram utilizadas em 10% dos casos para ajustar a categoria final.
- Mais presentes nas RM que nas TC
  ↑164 RM e 118 CT
  ↓118 RM e 73 CT
- Difícil discernir o impacto individual de cada característica auxiliar.
- Utilizadas em 9% dos casos para ajustar a categoria final de LI-RADS
  (119 RM e 82 TC)



Discussão

Superior quando comparado a 6 estudos anteriores:

- ✓ Múltiplos centros hospitalares, internacional, com uma “pool” diversificada de radiologistas;
- ✓ Maior número de radiologistas;
- ✓ Variedade de fontes de casos;
- ✓ Sem “módulos de treino”;
- ✓ Combinação de todas categorias de LI-RADS;
- ✓ Redução da variabilidade da condições de visualização das imagens;
- ✓ Sem necessidade de verificação patológica (eliminando o viés de verificação);
- ✓ Inclusão das características auxiliares e regras de “desempate”.

Discussão - Limitações

- Casos limitados a imagens selecionadas e anotadas;
- Impossibilidade de alterar as definições de janela;
- Imagens com qualidade superior à encontrada na prática de rotina;
- LR-5 e LR-5V agrupadas;
- Impossibilidade de tratar LR-M como categoria separada;
- Falta de “standard” para o presente estudo.

Discussão - Conclusões

LI-RADS revelou:

- **Nível de concordância substancial** (0,70-0,71) nas categorias LR-1 a LR-5 e LR-M;
- **Nível de concordância quase perfeita** (>0,8) nas caraterísticas major;
- Semelhante IRR para TC e RM;
- **Factores demográficos dos "leitores" afectam minimamente o CCI – LI-RADS pode ser usado por radiologistas de todos os backgrounds e níveis de experiência;**
- Características auxiliares só contribuíram para alterações de categoria em 10% dos casos e as regras de desempate em 9% dos casos;

↓

- Os autores sugerem futuras investigações sobre estes aspetos e a sua otimização na inclusão no algoritmo.