



HCC screening with ultrasound: assessment of quality using ultrasound LI-RADS score

Michael J. King¹ · Karen M. Lee¹ · Sonam Rosberger¹ · Hsin-hui Huang^{2,3} · Gabriela Hernandez Meza⁴ · Sara Lewis^{1,2} · Bachir Taouli^{1,2}

Received: 15 June 2022 / Revised: 4 October 2022 / Accepted: 5 October 2022 / Published online: 15 October 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Purpose To describe ultrasound (US) quality for hepatocellular carcinoma (HCC) screening/surveillance using the US LI-RADS scoring system, and to assess predictive factors of worse US quality scores.

Methods This retrospective study included adult patients ($n = 470$; M/F 264/206, median age 59y) at risk for HCC that underwent US for HCC screening/surveillance. US examinations were independently reviewed by 2 radiologists that assigned a visualization score (A: no/minimal, B: moderate, C: severe limitation) and US diagnostic category (US LI-RADS 1: negative, US LI-RADS 2: subthreshold, US LI-RADS 3: positive) to each study. A generalized linear mixed model was used to assess the predictive factors of worse visualization score using OR (odds ratio) statistics. Simple Kappa coefficient (K) assessed inter-reader agreement.

Results For readers 1 and 2, 295/320 (62.8%/68.1%) cases were scored A, 153/134 (32.6%/28.5%) were scored B, and 22/16 (4.6%/3.4%) were scored C, respectively. There was moderate inter-reader agreement for US LI-RADS visualization score ($K = 0.478$) and 100% concordance for US diagnostic category ($K = 1$), with 30 (6.4%) cases scored as positive (US LI-RADS 3). Cirrhosis and obesity were significant independent predictors of worse visualization scores (B/C) (cirrhosis: OR 10.4 confidence intervals: [4.25–25.48], $p < 0.001$; obesity: OR 3.61 [2.11–6.20], $p < 0.001$). Of the 30 lesions scored as US LI-RADS 3, 9 were characterized as probable or definite HCC on confirmatory CT/MRI, yielding a PPV of 30% (9/30) and a false-positive rate of 70% (21/30).

Conclusion Moderate to severe limitations in quality of US performed for HCC screening/surveillance was observed in approximately one-third of patients. Patients with cirrhosis and/or elevated BMI have poorer quality US studies and may benefit from other screening modalities such as CT or MRI.

Keywords Ultrasound · Hepatocellular carcinoma · Cancer screening

✉ Michael J. King
michael.j.king@mountsinai.org
Karen M. Lee
karen.lee@mountsinai.org
Sonam Rosberger
sonam.rosberger@mountsinai.org
Hsin-hui Huang
hsinhui.huang@gmail.com
Gabriela Hernandez Meza
gabriela.hernandezmeza@mountsinai.org
Sara Lewis
sara.lewis@mountsinai.org
Bachir Taouli
bachir.taouli@mountsinai.org

¹ Department of Diagnostic, Molecular and Interventional Radiology, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1234, New York, NY 10029-6574, USA

² BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

³ Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, USA

⁴ Department of Anesthesiology, Perioperative and Pain Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Abbreviations

AFP	Alpha-fetoprotein
ALD	Alcoholic liver disease
BMI	Body mass index
CT	Computed tomography
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HCC	Hepatocellular carcinoma
LI-RADS	Liver imaging reporting and data system
MRI	Magnetic resonance imaging
NASH	Nonalcoholic steatohepatitis
US	Ultrasound

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common cancer and second most common cause of cancer mortality worldwide, with a 5-year survival rate of 18% [1]. It is one of the leading causes of death in patients with cirrhosis, and early tumor detection in the form of biannual screening/surveillance has improved survival in these high-risk patients [2–6]. Ultrasound (US) is the most widely used modality for HCC screening/surveillance (with or without serum alpha-fetoprotein (AFP) testing) and is recommended as the preferred imaging test by all major liver societies worldwide [7–9].

The American College of Radiology (ACR) has developed the US Liver Imaging Reporting and Data System (US LI-RADS) algorithm to streamline such examinations, improve communication among physicians, and guide patient management utilizing a distinct visualization score and an assigned US category. The visualization score is a subjective assessment and informs the expected sensitivity of the study with three possibilities: score A (no or minimal limitations), score B (moderate limitations), and score C (severe limitations). The US category guides management and has three possibilities: US LI-RADS 1 (negative; no US evidence of HCC), US LI-RADS 2 (subthreshold; observation(s) < 10 mm in diameter, not definitely benign), and US LI-RADS 3 (positive; observation(s) ≥ 10 mm in diameter, not definitely benign, or new thrombus in vein) [7, 10]. These scores and categories apply to the patient level rather than observation level.

A variety of technical and patient-related factors may affect the quality of the US examination, with examples including suboptimal beam penetration related to the patient's body habitus or background liver heterogeneity, inconsistent or unsatisfactory scanning technique by the sonographer (when applicable), and inability of the patient to follow breathing instructions, among many others. Thus, patients undergoing HCC screening/surveillance with US

may be at-risk of having an early-stage HCC missed, potentially leading to increased morbidity and mortality.

Two studies have evaluated the clinical performance of US LI-RADS and assessed various predictors of visualization scores, including a large study of over 2000 patients [11, 12]. To complement these studies, we tested the effect of demographics and clinical factors on US visualization score. The purpose of our study is to describe US quality for HCC screening/surveillance using the US LI-RADS scoring system, and to assess predictive factors of worse US quality scores.

Materials and methods

Patients

This retrospective, Health Insurance Portability and Accountability Act (HIPAA)-compliant single-center study was approved by our institutional review board with exemption for patient consent. The Department of Radiology imaging database was queried for abdominal US examinations performed at our major urban academic tertiary transplant center between January 1, 2017 and December 31, 2017. Included patients were adults (age ≥ 18 years) deemed at-risk for HCC (those with non-cirrhotic chronic hepatitis B (HBV) or cirrhosis not due to congenital hepatic fibrosis or a vascular disorder) that underwent US for routine HCC screening/surveillance. The patient population consisted of 470 patients (M/F 264/206, median age 59.0 years, IQR 49.8–66.0, range 23–87); this population is a subset of the cohort published in a recent study assessing patterns of HCC screening at our institution [13].

Demographic and clinical data was recorded for each patient from the electronic medical record, including age, gender, race/ethnicity, BMI, presence/absence of cirrhosis, etiology of underlying liver disease, and Child–Pugh score (where applicable). Characteristics of our study cohort are summarized in Table 1.

Image acquisition

Abdominal US examinations were performed at outpatient imaging facilities at our hospital utilizing Philips EPIQ or Philips iU22 US systems (Philips Healthcare, Andover, MA). The US examinations were performed by one of 15 technologists over the course of the year according to a standardized institutional protocol.

Image analysis

Visualization scores were randomly assigned to the abdominal US examinations independently by two board-certified

Table 1 Characteristics of our study cohort ($n=470$)

Demographic or clinical characteristic	Value
Age (y)	59.0 [49.8–66.0] [#]
Gender(Male/Female)	264 (56.2%)/206 (43.8%)
Race	
Caucasian	232 (49.4%)
Black	53 (11.3%)
Asian	72 (15.3%)
Unknown	113 (24.0%)
Ethnicity	
Hispanic	153 (32.6%)
Non-Hispanic	251 (53.4%)
Unknown	66 (14.0%)
Body mass index (BMI) (kg/m ²)	27.0 [23.7–31.6] [#]
Cirrhosis present/absent	367 (78.1%)/103 (21.9%)
Liver disease etiology	
HCV	180 (38.3%)
HBV	133 (28.3%)
Alcohol	63 (13.4%)
NASH	42 (8.9%)
Other*	52 (11.1%)
Child–Pugh score**	
A	257 (70.0%)
B	72 (19.6%)
C	38 (10.4%)

Values are given as numerical value with percentages of each respective subgroup unless otherwise indicated

[#] Value given as median [interquartile range]

*Includes sarcoidosis ($n=1$), progressive familial intrahepatic cholestasis ($n=1$), biliary atresia ($n=2$), hemochromatosis ($n=2$), Budd-Chiari syndrome ($n=1$), primary biliary cholangitis ($n=15$), cryptogenic ($n=12$), auto-immune hepatitis ($n=12$), primary sclerosing cholangitis ($n=6$)

**only for patients with cirrhosis

fellowship-trained abdominal radiologists (KL, SR), with 16- and 20-years' experience, respectively, using PACS (Centricity 3.0, General Electric Medical Systems) [14]. Examples of score A include studies where the liver is homogeneous or minimally heterogeneous, there is minimal beam attenuation or shadowing, or where the liver is visualized in near entirety. Examples of score B include studies where the liver is moderately heterogeneous, there is moderate beam attenuation or shadowing, or some portions of the liver or diaphragm are not visualized. Examples of score C include studies where the liver is severely heterogeneous, there is severe beam attenuation or shadowing, or most of the liver and/or diaphragm are not visualized (Fig. 1) [10].

Statistical analysis

Demographic, clinical, and imaging data were summarized using descriptive statistics. The difference in baseline characteristics between cirrhotic and noncirrhotic patients was assessed using the Wilcoxon-rank sum and Fisher's exact tests. A generalized linear mixed model, accounting for the correlation of the measurements from the same patients and readers, was used to test the effects of cirrhosis and obesity (defined as BMI ≥ 30 kg/m²) on worse visualization score (B or C), after adjusting for covariates of age, gender, and race/ethnicity utilizing odds ratio (OR) statistics. Kappa coefficient (K) assessed inter-reader visualization score and diagnostic category agreement using the following categories: 0 to 0.20 as slight, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1 as almost perfect agreement [15]. A p -value ≤ 0.05 represented an acceptable level of statistical significance in our study. All statistical tests were conducted using R statistical software (version 4.1.3, Vienna, Austria).

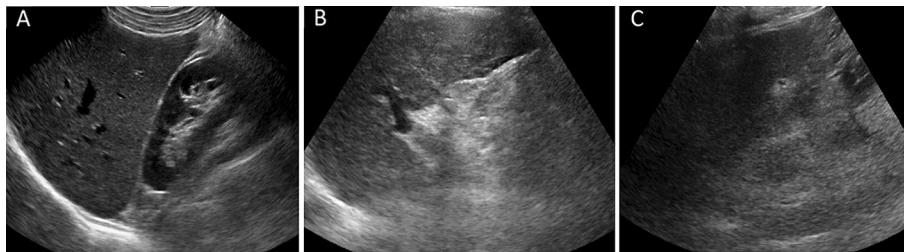


Fig. 1 Illustration of ultrasound (US) LI-RADS visualization scores in patients undergoing US for HCC screening. **A** US LI-RADS visualization score A (no or minimal limitations): 42-year-old male with BMI of 28.6 kg/m² and noncirrhotic chronic HBV. Longitudinal grayscale US image of right hepatic lobe demonstrates noncirrhotic, homogeneous liver parenchyma with no beam attenuation or shadowing. Portion of liver imaged is visualized in its entirety with clearly visible right hemidiaphragm and liver border. No observation identified. **B** US LI-RADS visualization score B (moderate limitations): 55-year-old female with BMI of 29.6 kg/m² and HCV cirrhosis. Lon-

gitudinal grayscale image of right hepatic lobe demonstrates moderately heterogeneous, cirrhotic liver parenchyma with some beam attenuation and shadowing, with a portion of liver and right hemidiaphragm not visualized. No focal observation clearly identified. **C** US LI-RADS visualization score C (severe limitations): 62-year-old male with BMI of 32.5 kg/m² and NASH cirrhosis. Longitudinal grayscale image of right hepatic lobe demonstrates severely heterogeneous, cirrhotic liver parenchyma with significant beam attenuation and shadowing, with majority of right lobe and right hemidiaphragm not visualized. No focal observation clearly identified

Results

US findings

For reader 1, 295 (62.8%) cases were scored as visualization score A, 153 (32.6%) as score B, and 22 (4.6%) as score C. For reader 2, 320 (68.1%) cases were scored as visualization score A, 134 (28.5%) as score B, and 16 (3.4%) as score C (Fig. 2). There was moderate inter-reader agreement for US LI-RADS visualization score ($K=0.478$) and there was 100% concordance among readers for assigned category (Kappa = 1), with 430 (91.5%) cases scored as US LI-RADS 1, 10 (2.1%) as US LI-RADS 2, and 30 (6.4%) as US LI-RADS 3. Of the 30 cases scored LI-RADS 3 by both readers, 9 were ultimately characterized as probable (LI-RADS 4) or definite (LI-RADS 5) HCC on follow-up confirmatory CT or MRI, yielding a positive predictive value (PPV) of 30% (9/30), with 21 false-positives ultimately characterized as LI-RADS 1 ($n=12$), LI-RADS 2 ($n=3$), and LI-RADS 3 ($n=6$) on

follow-up cross-sectional imaging (Figs. 3 and 4, Table 2). There were no cases that had more than 1 lesion.

Predictive factors of LI-RADS visualization score

The interaction between cirrhosis and BMI was not significant (OR 0.96 [0.83–1.10], $p=0.53$), and there was no significant difference in the baseline characteristics of age, race/ethnicity, BMI, and etiology of liver disease between cirrhotic and noncirrhotic patients (all p -values <0.001). Cirrhosis (OR 10.4 [4.25–25.48], $p<0.001$) and obesity (OR 3.61 [2.11–6.20], $p<0.001$) predicted worse visualization scores (B or C) after adjusting for covariates of age, gender and race/ethnicity.

Discussion

In this study performed in patients undergoing HCC screening/surveillance, we demonstrated acceptable inter-reader agreement for US LI-RADS visualization score ($K=0.478$), with approximately one-third of cases being moderately

Fig. 2 Ultrasound LI-RADS visualization score breakdown per reader in our cohort ($n=470$)

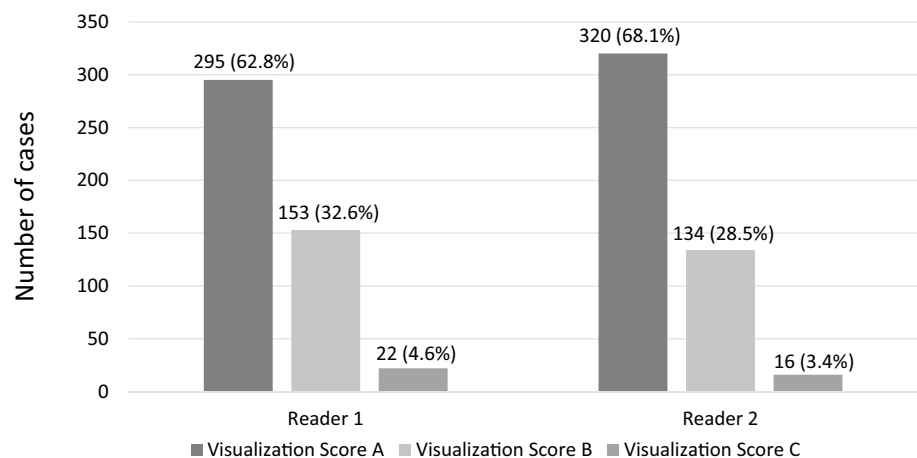


Fig. 3 Illustration of US LI-RADS chronic HCV cirrhosis. Transverse grayscale image of left hepatic lobe demonstrates a 3.9 cm homogeneously echogenic subcapsular observation in segment 2 (A; arrow). This was confirmed to be hepatocellular carcinoma on sub-

sequent MRI (LI-RADS 5) (B, C; arrows), characterized by arterial phase hyperenhancement, portal venous washout, and enhancing capsule

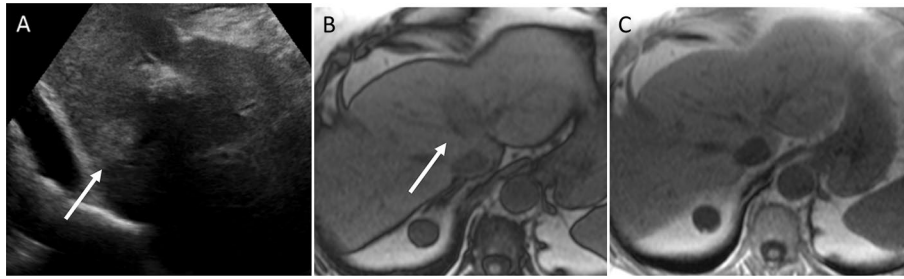


Fig. 4 Illustration of US LI-RADS 3 false-positive. **A** US-3 (false-positive): 50-year-old female with noncirrhotic chronic HBV. Longitudinal grayscale image of right hepatic lobe demonstrates a 2.2 cm homogeneously echogenic subcapsular observation near the hepatic

dome (arrow). **B, C.** Subsequent confirmatory MRI demonstrates a focal region of signal intensity loss on opposed-phase T1 weighted image compared to in-phase image (arrow) without abnormal enhancement (not shown), consistent with focal fat deposition

diagnostic (visualization score B) and approximately 5% of cases being essentially nondiagnostic (visualization score C). Of the 30 US LI-RADS 3 lesions requiring confirmatory cross-sectional imaging, 9 were ultimately characterized as probable or definite HCC on follow-up CT or MRI, yielding a PPV of 30% (9/30). We also demonstrated that cirrhosis and obesity are both similar independent predictors of worse US quality, revealing the high impact cirrhosis and obesity have on worse visualization score. Our 100% concordance among readers in assigning US LI-RADS category is not surprising and is likely related to sonographer annotations, however, this may not be the case elsewhere.

These results are closely aligned with the largest study published to date assessing US LI-RADS clinical performance in over 2000 patients, where the authors demonstrated a visualization score breakdown (76.8% A, 18.9% B, 4.2% C) and US category breakdown (90.4% US LI-RADS 1, 4.6% US LI-RADS 2, 4.9% US LI-RADS 3) similar to our study [11]. Our moderate inter-reader agreement for visualization score aligns closely with a very recent Canadian study by Kiri et al. assessing US LI-RADS visualization score in 237 US examinations that also demonstrated moderate inter-reader agreement ($K=0.51$) [12]. To complement these published studies, we tested additional factors that may affect US quality, including race/ethnicity and BMI. Like in our results, the study by Millet et al. demonstrated a PPV of 35% for US LI-RADS; however, it demonstrated high sensitivity of 82% for HCC detection despite other reported sensitivities as low as 45% in a recent meta-analysis [3, 11]. Our results closely align with a study assessing outcomes of 267 US LI-RADS 3 observations; in which the authors demonstrated a PPV of a US LI-RADS 3 observation for probable or definite HCC of 18.8% [16]. No follow-up was sought on our patients with US LI-RADS 1 or 2 scores, preventing us from calculating sensitivity and specificity.

Increased BMI is widely associated with poorer quality US, both scientifically and in clinical practice. In the aforementioned Canadian study, the authors demonstrated similar

results in that patients with $BMI \geq 25 \text{ kg/m}^2$ were more likely to have visualization scores B or C at uni- and multi-variable analysis [12]. In another recent study of 352 patients assessing predictors of US failure to detect HCC, the authors demonstrated significantly reduced US sensitivity relative to CT or MRI in patients with $BMI \geq 30 \text{ kg/m}^2$ or NASH [17]. A study of 941 patients undergoing US HCC screening demonstrated that US was inadequate for excluding HCC in over one-third of patients with $BMI > 35 \text{ kg/m}^2$, Child–Pugh C cirrhosis, or NASH cirrhosis [18]. Anecdotally, a published survey of 14 sonographers revealed a general consensus that the best abdominal US images are obtained in patients with normal BMI, whereas images obtained from patients with $BMI 30.0$ to 34.9 kg/m^2 were limited in quality [19]. These findings are due to the fact that the mean depth of insonation is deeper for obese patients requiring a greater distance that the US waves have to travel, resulting in greater absorption and dispersion in the surrounding tissues with increased refraction backscatter and decreased signal-to-noise ratio (SNR) [20]. Decreased SNR leads to impaired US visualization of liver lesions and reduces the sensitivity of US screening/surveillance; combined with recently published data, these findings overall suggest that overweight or obese patients undergoing HCC screening/surveillance with US may benefit from other modalities to best detect potentially curable HCC, to allow early application of curative therapies, and to reduce morbidity and/or mortality.

HCC screening is performed with CT or MRI in many centers given the improved sensitivity and ability to detect early-stage tumors associated with these modalities [13, 21]. In a prospective Korean study directly comparing the performance of US and MRI for HCC screening/surveillance, the PPVs for HCC detection with US and MRI were 17% and 54%, respectively, and sensitivities for HCC detection were 28% and 86%, respectively [22]. The higher cost of these modalities, along with increased radiation exposure (CT) and relatively long exam times (MRI), raise questions about which imaging modality is optimal for

Table 2 Details of the 40 cases scored as either US LI-RADS 2 or US LI-RADS 3 by both readers

Case	Cirrhosis	Lesion size (cm)	Lesion echogenicity	TIV	US LI-RADS Score	CT/MR LI-RADS Score
1	No	0.2	Hyperechoic	No	2	
2	Yes	0.7	Hypoechoic	No	2	
3	No	0.9	Hyperechoic	No	2	
4	No	0.5	Hyperechoic	No	2	
5	No	0.4	Hyperechoic	No	2	
6	No	0.9	Hyperechoic	No	2	
7	Yes	0.9	Hypoechoic	No	2	
8	No	0.8	Hyperechoic	No	2	
9	No	0.8	Hyperechoic	No	2	
10	Yes	0.9	Hyperechoic	No	2	
11	Yes	1.9	Hyperechoic	No	3	1 (hemorrhagic cyst)
12	Yes	N/A	N/A	Yes	3	1 (bland portal vein thrombus)
13	Yes	2.2	Hyperechoic	No	3	1 (hemangioma)
14	Yes	N/A	N/A	Yes	3	5 (tumor thrombus, HCC see on MRI, not seen on US)
15	Yes	1.9	Hyperechoic	No	3	5 (HCC)
16	Yes	3.4	Hyperechoic	No	3	5 (HCC)
17	Yes	2.8	Hyperechoic	No	3	5 (HCC)
18	Yes	1.6	Hypoechoic	No	3	5 (HCC)
19	Yes	3.9	Hyperechoic	No	3	1 (hemangioma)
20	Yes	2.1	Hyperechoic	No	3	2 (sclerosing hemangioma)
21	No	1.2	Hypoechoic	No	3	1 (no MRI correlate to US finding)
22	Yes	1.5	Hyperechoic	No	3	4 (probable HCC)
23	Yes	2.2	Hyperechoic	No	3	3 (regenerative/dysplastic nodule)
24	No	1.2	Hypoechoic	No	3	1 (no MRI correlate to US finding)
25	Yes	1	Hypoechoic	No	3	3 (regenerative/dysplastic nodule)
26	Yes	N/A	N/A	Yes	3	5 (infiltrative HCC with TIV)
27	Yes	3.5	Hyperechoic	No	3	3 (regenerative/dysplastic nodule)
28	Yes	1.1	Isoechoic	No	3	1 (no MRI correlate to US finding)
29	Yes	2.4	Hypoechoic	No	3	3 (regenerative/dysplastic nodule)
30	Yes	1	Hyperechoic	No	3	3 (regenerative/dysplastic nodule)
31	Yes	2	Hypoechoic	No	3	1 (no MRI correlate to US finding)
32	Yes	3.2	Hyperechoic	No	3	4 (probable HCC)
33	Yes	N/A	N/A	Yes	3	1 (bland portal vein thrombus)
34	Yes	1.1	Hyperechoic	No	3	4 (probable HCC)
35	Yes	1.7	Hyperechoic	Yes	3	2 (hemangioma)
36	No	2.6	Hyperechoic	No	3	2 (hemangioma)
37	Yes	1.3	Hyperechoic	No	3	3 (regenerative/dysplastic nodule)
38	Yes	1.1	Hyperechoic	No	3	1 (no MRI correlate to US finding)
39	Yes	1.7	Hypoechoic	No	3	1 (no MRI correlate to US finding)
40	Yes	2.2	Hyperechoic	No	3	1 (focal fat deposition)

HCC screening/surveillance. These drawbacks must be reconciled with the fact that US screening generally demonstrates higher false-positive rates than CT or MRI, often revealing indeterminate findings that require additional imaging and/or biopsy, adding to additional potential cost and harm [21, 23, 24].

In our study, a substantial proportion of patients (37% and 32%, for readers 1 and 2, respectively) were assigned visualization scores of B or C, highlighting the suboptimal US image quality frequently encountered in patients typically undergoing HCC screening/surveillance. In a very recent study of 3169 patients who underwent multiple

HCC surveillance US examinations, the probability of a repeat visualization score C on subsequent US increased with a higher number of preceding examinations with score C [25]. Potential future screening directions for this population include a patient-centered quality scoring system incorporating individual patient factors (such as presence of cirrhosis and BMI) to adequately stratify risk and identify those in whom US sensitivity may be reduced; in these patients, and especially in those with previous US examinations with score C, screening with CT or MRI would likely be beneficial. Abbreviated MRI techniques could also be used to mitigate the financial and time cost of a complete MRI, while providing more comprehensive information than US about the liver and other abdominal organs potentially relevant to transplant candidates [26–29].

Our study has several limitations, most notably being its retrospective, single-center design. Our patient population being from a densely populated, urban setting is not necessarily reflective of the true national or global human population. Although patients were scanned according to a standardized institutional protocol, obtaining images from 15 different sonographers and 2 different US systems from the same vendor may have introduced variability in image quality. Our limited sample size, limited number of studies categorized as US LI-RADS 3, and suboptimal distribution of clinical variables prevented us from controlling for too many covariates (including etiology of underlying liver disease and Child–Pugh score) in the statistical analysis, and from providing more robust descriptive statistics. Finally, no follow-up was sought on any patients in our cohort receiving US LI-RADS 1 or US LI-RADS 2 scores, preventing us from calculating sensitivity and specificity.

In conclusion, we report an acceptable inter-reader agreement when assigning a US LI-RADS visualization score and 100% concordance in assigning a diagnostic category. Moderate to severe limitations in quality of US performed for HCC screening/surveillance was observed in approximately a third of patients in our cohort. Patients with cirrhosis and/or obesity are more likely to have poorer quality US studies and may benefit from other screening modalities such as CT or MRI. Future research should consider stratifying patients at-risk into adequate screening tests.

Acknowledgements The remaining authors did not receive support from any organization for the submitted work.

Author contribution Bachir Taouli, MD, MHA: Research support/grant: Bayer, Takeda, Regeneron, Helio Health, Siemens, Echoscens. Consultant: Bayer, Guerbet.

Funding The remaining authors have no relevant financial or non-financial interests to disclose.

References

1. Tang, A., et al., *Epidemiology of hepatocellular carcinoma: target population for surveillance and diagnosis*. *Abdom Radiol (NY)*, 2018. **43**(1): p. 13–25.
2. Jemal, A., et al., *Annual Report to the Nation on the Status of Cancer, 1975–2014, Featuring Survival*. *J Natl Cancer Inst*, 2017. **109**(9).
3. Tzartzeva, K., et al., *Surveillance Imaging and Alpha Feto-protein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis*. *Gastroenterology*, 2018. **154**(6): p. 1706–1718.e1.
4. Singal, A.G., A. Pillai, and J. Tiro, *Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis*. *PLoS Med*, 2014. **11**(4): p. e1001624.
5. Kansagara, D., et al., *Screening for hepatocellular carcinoma in chronic liver disease: a systematic review*. *Ann Intern Med*, 2014. **161**(4): p. 261–9.
6. Zhang, B.H., B.H. Yang, and Z.Y. Tang, *Randomized controlled trial of screening for hepatocellular carcinoma*. *J Cancer Res Clin Oncol*, 2004. **130**(7): p. 417–22.
7. Morgan, T.A., et al., *US LI-RADS: ultrasound liver imaging reporting and data system for screening and surveillance of hepatocellular carcinoma*. *Abdom Radiol (NY)*, 2018. **43**(1): p. 41–55.
8. Song, P., et al., *The clinical management of hepatocellular carcinoma worldwide: A concise review and comparison of current guidelines from 2001 to 2017*. *Biosci Trends*, 2017. **11**(4): p. 389–398.
9. Fetzer, D.T., et al., *Screening and Surveillance of Hepatocellular Carcinoma: An Introduction to Ultrasound Liver Imaging Reporting and Data System*. *Radiol Clin North Am*, 2017. **55**(6): p. 1197–1209.
10. Radiology, A.C.o. *Ultrasound LI-RADS v2017*. [cited 2022 January 19]; Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/Ultrasound-LI-RADS-v2017>.
11. Millet, J.D., et al., *ACR Ultrasound Liver Reporting and Data System: Multicenter Assessment of Clinical Performance at One Year*. *J Am Coll Radiol*, 2019. **16**(12): p. 1656–1662.
12. Kiri, L., et al., *US LI-RADS Visualization Score: Interobserver Variability and Association With Cause of Liver Disease, Sex, and Body Mass Index*. *Can Assoc Radiol J*, 2022. **73**(1): p. 68–74.
13. Hernandez-Meza, G., et al., *MRI is the most commonly used imaging modality for HCC screening at a tertiary care transplant center*. *Abdom Radiol (NY)*, 2021. **46**(11): p. 5142–5151.
14. Radiology, A.C.o. *Ultrasound LI-RADS® v2017*. 2017 06/01/2021]; Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/Ultrasound-LI-RADS-v2017>.
15. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data*. *Biometrics*, 1977. **33**(1): p. 159–74.
16. Sevco, T.J., et al., *Ultrasound (US) LI-RADS: Outcomes of Category US-3 Observations*. *AJR Am J Roentgenol*, 2021. **217**(3): p. 644–650.
17. Samoylova, M.L., et al., *Predictors of Ultrasound Failure to Detect Hepatocellular Carcinoma*. *Liver Transpl*, 2018. **24**(9): p. 1171–1177.
18. Simmons, O., et al., *Predictors of adequate ultrasound quality for hepatocellular carcinoma surveillance in patients with cirrhosis*. *Aliment Pharmacol Ther*, 2017. **45**(1): p. 169–177.
19. Brahee, D.D., et al., *Body Mass Index and Abdominal Ultrasound Image Quality: A Pilot Survey of Sonographers*. *Journal of Diagnostic Medical Sonography*, 2013. **29**(2): p. 66–72.

20. Paladini, D., *Sonography in obese and overweight pregnant women: clinical, medicolegal and technical issues*. *Ultrasound Obstet Gynecol*, 2009. **33**(6): p. 720-9.
21. Joshi, K., et al., *Hepatocellular carcinoma surveillance: a national survey of current practices in the USA*. *Dig Dis Sci*, 2014. **59**(12): p. 3073-7.
22. Kim, S.Y., et al., *MRI With Liver-Specific Contrast for Surveillance of Patients With Cirrhosis at High Risk of Hepatocellular Carcinoma*. *JAMA Oncol*, 2017. **3**(4): p. 456-463.
23. Colli, A., et al., *Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review*. *Am J Gastroenterol*, 2006. **101**(3): p. 513-23.
24. Arguedas, M.R., et al., *Screening for hepatocellular carcinoma in patients with hepatitis C cirrhosis: a cost-utility analysis*. *Am J Gastroenterol*, 2003. **98**(3): p. 679-90.
25. Tiyyarattanachai, T., D.T. Fetzer, and A. Kamaya, *Multicenter Study of ACR Ultrasound LI-RADS Visualization Scores on Serial Examinations: Implications for Changes in Surveillance Strategies*. *AJR Am J Roentgenol*, 2022.
26. An, J.Y., et al., *Abbreviated MRI for Hepatocellular Carcinoma Screening and Surveillance*. *Radiographics*, 2020. **40**(7): p. 1916-1931.
27. Bartolotta, T.V., et al., *Contrast-enhanced ultrasound of hepatocellular carcinoma: where do we stand?* *Ultrasonography*, 2019. **38**(3): p. 200-214.
28. Besa, C., et al., *Hepatocellular carcinoma detection: diagnostic performance of a simulated abbreviated MRI protocol combining diffusion-weighted and T1-weighted imaging at the delayed phase post gadoxetic acid*. *Abdominal Radiology*, 2017. **42**(1): p. 179-190.
29. Vietti Violi, N., et al., *Abbreviated Magnetic Resonance Imaging for HCC Surveillance*. *Clin Liver Dis (Hoboken)*, 2021. **17**(3): p. 133-138.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.