Development of a Screening Liver Ultrasound Grading Criteria to Establish the Ability to Subjectively Identify Liver Lesions
• Authors
  • Emily Johnson, MD
  • Beverly Hasimoto, MD
  • Dawna Kramer, MD
• Disclosures
  • None
Purpose

• To establish the need for **standardized reporting** for screening liver ultrasound (US)

• Develop a **grading scale** for the ability to subjectively detect liver lesions on US
Background

- Patients with risk factors for hepatocellular carcinoma are regularly screened to identify suspicious liver lesions using US
- US is safe, fast, and low cost compared to CT or MRI
- Technical limitations of US include:
  - Poor penetration
  - Liver heterogeneity
  - Anatomic constraints, e.g., Liver high in the chest, morbid obesity
- Decreased US sensitivity may prompt the use of alternate screening modalities for future screening
Method

• Retrospective cohort study
  • All dictations of screening liver US studies, 189 cases
  • January 2015 – August 2015
  • Primary objective - identify explicit reference to US quality/ability to resolve a lesion in the final report
    • This does not include general characteristics of the liver that may reduce the ability to see lesions such as increased echogenicity or heterogeneity.
  • Secondary objective
    • Frequency of suspicious liver lesions
    • Frequency of incidental findings requiring follow up

• Development of a 3-point scale for US ability to resolve a liver lesion
  • Consensus review by 2 experienced US attendings and lead sonographer
  • Created from small series of representative liver US exams
Method

- Evaluation of reproducibility of screening liver US grading criteria
- Retrospective review of 80 screening liver US cases
  - Reviewed by 3 independent graders and assigned a score
  - Static and cine US images
  - Blinded to patient clinical history and prior exams
  - Calculated rate of inter-observer agreement amongst the 3 graders
- Department wide roll out after ensuring satisfactory reproducibility
  - Grading scale
  - Standardized dictation template
  - Attending and sonographer education
Results

• Screening liver US dictation review, 189 studies
  • 11 (5.8%) with explicit reference to US quality or ability to resolve a lesion
  • Secondary observations
    • 3 (1.6%) identified a suspicious liver lesion
    • 5 (2.6%) identified an incidental finding needing follow up
      – 3 gallbladder polyps, 1 adenomyomatosis vs air, 1 adenomyomatosis vs mass

• Inter-observer agreement, 80 studies
  • 63.7-67.5% agreement between any pair of observers
  • No cases were graded in the opposite manner
    • E.g. (1 “good” and 3 “poor”)
  • Grade 1 “good” – 39.91% of cases
  • Grade 2 “fair” – 50.17%
  • Grade 3 “poor” – 9.92%
Liver Screening Categorization:
The subjective ability to identify a focal liver lesion

**Good – Grade 1**
- Homogeneous echotexture
- Normal echogenicity
- Diaphragm well seen posteriorly
- Sharp vessel margins

**Fair – Grade 2**
- Mildly heterogeneous echotexture
- Mildly echogenicity
- Diaphragm partially seen posteriorly
- Vessel margins slightly obscured

**Poor – Grade 3**
- Very echogenic
- Very heterogeneous echotexture
- Poor penetration - diaphragm not seen posteriorly
- Poor resolution - vessels obscured
- Liver positioned high in thorax
Liver:

Size [ ]
Echogenicity [ ]
Surface contour [ ]
Perceived sensitivity of ultrasound for detecting focal lesions [ ]
Focal lesions [ ]

If perceived sensitivity of ultrasound is compromised, an alternative imaging modality for screening should be considered.

Impression:

[ ]

Liver Appearance Sensitivity Definition:
Good:
Normal echogenicity
Diaphragm well visualized
Vessel margins sharp

Fair:
Mildly heterogeneous
Mildly echogenic
Diaphragm partially visualized
Vessel slightly obscured

Compromised:
Very echogenic
Very heterogeneous
Poor penetration—diaphragm not seen
Poor resolution—vessels obscured
Liver high in thorax
Discussion

- Review of dictations for screening liver US revealed very low acknowledgement of quality in evaluation of the liver for lesions.
  - Only 5.8% of cases commented on quality (either good or bad) and subsequent analysis revealed nearly 10% of cases were classified as “poor” for ability to detect a liver lesion.
  - Discrepancy suggests that the hepatologist were not being alerted that an alternate screening method (CT or MRI) may be beneficial due to probable low US sensitivity
  - Hepatology was very interested in a direct way of communicating utility of US as a screening modality for patients at risk of HCC- not having to “read between the lines” in the text of the report.

- Grading scale remains a subjective assessment that should be determined by the sonographer and the radiologist

- Anecdotally, the inter-observer agreement increased with the increasing number of cases reviewed (increased agreement in the second half of the cases reviewed compared to the first half), which is not unexpected as there was increasing familiarity with the criteria

- Our rate of inter-observer agreement was felt to be acceptable to proceed to implementing the grading scale for all screening liver US in the department after introduction to the attendings and sonographers with sample cases.
Additional Consideration

Our grading scale was developed prior to the release of the CEUS LI-RADS v2016, but generally reflects the cited limitations of US for evaluation of liver lesions.

Potential Pitfalls and Challenges

- Observation/nodule dimension less than 10mm
- Subdiaphragmatic or deep location
- Large body habitus
- Hepatic steatosis
- Very coarse heterogeneous cirrhotic liver
- Poorly cooperating patients
- Interfering bowel or gastric gas
- Nonlinear propagation artifact, a CEUS phenomenon, is associated with pseudo-enhancement following microbubble contrast injection. It is caused by the nonlinear propagation of sound through intervening microbubble-perfused tissue and is, therefore, most marked deep in the field of view.

ACR CEUS LI-RADS 2016; www.acr.org/Quality-Safety/Resources/LIRADS
Conclusion

• We established a **need for an explicit reference** to the ability to identify a lesion in screening liver US exams

• We created a **grading scale** to help quantify the ability to identify a lesion on screening liver US exams with reasonable inter-observer agreement

• Only a small number of screening liver US exams were graded as “poor” in the ability to identify a liver lesion
  • Alternate methods of screening should be considered for future follow up