Liver biopsy assessment of steatosis

Amar Paul Dhillon
Royal Free Hospital
RIBA, London
Thursday 28th November 2013
NAFLD didn’t exist before 2001 and liver biopsy assessment of steatosis was practically unnecessary.

Before 2001, critical assessment of steatosis was of little importance. Alcoholic steatohepatitis was the main thing.

Sheila Sherlock (1918-2001):
“There are two kinds of alcoholics: those who admit it; and those who don’t”

ie the whole of the ASH/NASH debate is confounded by the unreliability of the alcohol drinking clinical history.
The importance of NAFLD increased after 2001: we should be bothered about steatosis assessment because a LOT of people could be affected

- **Armstrong MJ et al. J Hepatology 56,234;2012**
  - NAFLD is the commonest cause of incidental LFT abnormalities in primary care (26.4%).

- **Bedogni G et al. Hepatology 42,44;2005**
  - Prevalence assessed in the general population (Modena) 18-75y with suspected liver disease on the basis of LFTs & hepatitis viral serology; and those without suspected liver disease.
  - NAFLD diagnosed by ultrasonography, and alcohol intake was assessed by using a 7-day diary.
    - The prevalence of NAFLD is:
      - 20% in people NOT suspected of having liver disease.
      - 25% in people suspected of having liver disease.
Is it safe to have “simple steatosis”?

  - 4% of patients with histologically determined “steatosis only” progress to cirrhosis over time (vs 22% with steatohepatitis; follow-up 8.3 +/- 5.4 years).
- There is a greater risk of steatohepatitis with increasing steatosis.
  - Chalasani N et al. J Hepatol 48,829;2008: “Compared to liver biopsies with mild steatosis, the odds of having definite steatohepatitis were 1.7 times greater among those with moderate, and 1.6 times greater among those with severe steatosis.”
- Liver biopsy cannot be used for population screening, obviously.
  - In the first instance ultrasound (& LFTs) are used to select patients with fatty liver for further investigation.
  - Machado MV & Cortez-Pinto H. J Hepatology 58,1007;2013:
    - “Ultrasound should be the first method to be used in a clinical setting.”
  - At some point we must clarify how liver biopsy fat assessment corresponds with imaging screening methods.
Is it safe to have “simple steatosis”? 

- Pais R et al. J Hepatol 59,550;2013
  - 70 patients with untreated NAFLD and with two biopsies performed more than one year apart.
  - Initially 25 patients had NAFL and 45 had NASH and/or advanced fibrosis.
  - 16/25 (64%) NAFL patients developed NASH (mean follow-up 3.7 +/- 2.1 years)
    - Average % steatosis (Kleiner system) in patients who progressed to steatohepatitis was 56% vs 34% in those who did not.
  - 25/70 (36%) patients had disease progression.
    - “These patients were not different from the others as far as clinical, biological, and metabolic characteristics, but had a significantly higher amount of steatosis (median 60% vs. 40%, p = 0.008).”
    - “In univariate analysis, only the amount of steatosis on initial biopsy correlated with disease progression (r = 0.325, p = 0.005).”
    - “In multivariate analysis, the amount of steatosis was the only independent factor associated with disease progression, even after adjustment for age, sex, BMI, and aminotransferase levels (r = 0.318, p = 0.01).”
Liver biopsy assessment of steatosis: should we be bothered?

- But the distinction of NASH from “simple” fatty liver still requires liver biopsy.
  - So you could say liver biopsy assessment of steatosis per se doesn’t matter since steatohepatitis is the most important thing, not steatosis.
- But it is still important to assess liver biopsy steatosis properly so that the non-invasive methods for the diagnosis of fatty liver are calibrated correctly.
- Also, steatosis is worth 3 of 8 points in the Kleiner NAS score which is currently the standard system for clinical trials and publications.
- Also, surgeons get excited by histological steatosis assessment for the selection of donor livers and planning hepatic resections.
Grading fat according to % of steatotic hepatocytes vs tissue sectional fat % area
% “of hepatocytes” vs % “of parenchymal involvement by steatosis”

  - “percentage (of) the surface area of a histological section involved.”
  - Steatosis: “grade 0<5% (“of hepatocytes”); grade 1 5–33%; grade 2 34–66%; grade 3 >67%.”
  - “percentage of liver parenchyma occupied by steatotic hepatocytes”
  - “the amount of surface area of parenchyma visually determined to be involved by steatosis.”
  - “Low- to medium-power evaluation of parenchymal involvement by steatosis: <30%, 30-60; >60.”
  - “divide the involved parenchyma by thirds, for example, 0–33%, 33%–66%, >66%”.
  - “Macrovesicular steatosis was graded 0–3 based on percent of hepatocytes in the biopsy involved (0 is none; 1 is up to 33%; 2 is 33–66%; 3 is >66%)”.


Grading fat according to % of steatotic hepatocytes vs tissue sectional fat % area

- Grading fat according to % of hepatocytes that are steatotic is an important issue that deserves attention because many hepatopathologists believe that is what they are doing.
- Brunt et al. Am J Gastroenterol. 94,2467;1999:
  - “Macrovesicular steatosis was graded 0–3 based on percent of hepatocytes in the biopsy involved”.
- However:
  - At low magnification, individual hepatocytes are not being counted.
  - Even if the observer thinks they are making % steatotic hepatocytes evaluation, they are not really.
  - At best, they might be subconsciously converting what is essentially a % fat area assessment into a % steatotic hepatocytes estimate based on their ideas of how many hepatocytes are contained within each of their microscopic fields of view.
  - The current state of the fat estimation art is as stated (Brunt et al.) in MacSween’s Pathology of the Liver (6th Ed; p326): ‘common semi-quantitative assessments for steatosis are based on the percentage (of) the surface area of a histological section involved’.
- Whatever liver steatosis screening imaging methods are counting, it is not % of steatotic hepatocytes.
The 0-100% steatosis scale is wrong for fat % area: the actual scale is 0-~50%
Liver biopsy assessment of steatosis:
what is the correct scale for the calibration of imaging methods?
NB: imaging does not count % steatotic hepatocytes

MRI vs subjective assessment
vs image analysis ("custom algorithm")
Morphometric, biochemical, and visual measurements of macrovesicular steatosis
Li M et al. Hum Pathol 42,356;2011

NB: biochemistry does not count % steatotic hepatocytes

- Twenty-six fresh liver (human deceased donors) specimens.
- Biochemical fat measurement.
  - fat content was defined as the weight percentage of fat.
- Formalin fixed tissue; H&E slides for subjective fat estimation.
  - % hepatocytes involved by macrovesicular steatosis.
- Digital images of slides were analyzed by computer morphometry.
  - fat defined content as the percentage of area occupied by fat droplets.
- The range of fat content in the specimens was:
  - 2.2-15% by biochemistry.
  - 0.8-82.5% by subjective visual estimation.
  - 0.3-19.6% by image analysis.
- “Visual estimation appeared to have a systematic bias, giving results nearly 4-fold higher than other methods.”
- “The absolute results of morphometry were essentially same as those of biochemical method, suggesting there was minimal systematic bias with morphometry.”
“Quantifying hepatic steatosis – more than meets the eye”
Levene A et al. Histopathol 60,971;2012

- Mouse model of NAFLD, and human liver biopsies.
- % steatosis and fat droplet size assessed in H&E and ORO stained sections by light microscopy and digital image analysis (DIA).
  - The percentage of hepatocytes involved by steatosis was assessed by counting hepatocytes at x400 magnification:
    - In six randomly selected fields (mouse tissue).
    - By counting every hepatocyte in the biopsy (human tissue).
  - “calculating the percentage of hepatocytes involved by steatosis using ORO staining found that 90–100% of hepatocytes in all the study groups contained steatosis, so this method could not be used”
    - This begs the question of what is meant by the term “steatotic hepatocyte”.
    - Which in turn raises the issue of how much fat a hepatocyte should contain to be considered an abnormal “steatotic hepatocyte”.
  - For DIA the % of steatosis by area in the liver was calculated using H&E or ORO-stained sections of frozen liver.
- ORO DIA is the most accurate method for detecting and quantifying steatosis as it showed the best correlation with triglyceride content in mouse and human liver.
Liver biopsy assessment of steatosis: “welcome to my world” example

• “Ultrasound shows fatty liver. NAFL(D)? NASH?”
• How much fat is there?
• Well, its kind of mildish; but then is it beginning to be moderate?
• And you “know” that mild is <33% and moderate is >33%; so you think: I’ll call it 30%.
• But then you think it doesn’t look anything like 30%.
• And that’s when the trouble starts.

<table>
<thead>
<tr>
<th>mFFA (%)</th>
<th>st eFFA* median (range)</th>
<th>st steatosis grade† (n = 12)</th>
<th>st</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3% (1–5%)</td>
<td>Normal 83% Mild 17%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10% (5–15%)</td>
<td>Normal 8% Mild 92%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10% (10–20%)</td>
<td>Mild 100%</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>25% (10–40%)</td>
<td>Mild 42% Moderate 58%</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>35% (20–40%)</td>
<td>Mild 42% Moderate 58%</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>60% (40–70%)</td>
<td>Moderate 75% Severe 25%</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>60% (35–70%)</td>
<td>Moderate 66% Severe 33%</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>75% (60–80%)</td>
<td>Moderate 17% Severe 83%</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>70% (50–85%)</td>
<td>Moderate 17% Severe 83%</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>80% (70–80%)</td>
<td>Moderate 8% Severe 92%</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>80% (70–90%)</td>
<td>Severe 100%</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>90% (80–95%)</td>
<td>Severe 100%</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>90% (80–100%)</td>
<td>Severe 100%</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>90% (90–100%)</td>
<td>Severe 100%</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>95% (90–100%)</td>
<td>Severe 100%</td>
<td></td>
</tr>
</tbody>
</table>

*Estimated fat percentage of the st based on area of parenchyma occupied by fat.
†Proportion of hepatopathologists (n = 12) assigning each grade of steatosis (normal/mild/moderate/severe).

The use of 33% and 66% cutoffs for mild vs moderate vs severe steatosis is arbitrary. Int liver pathology study group cutoffs (mFPA): normal (none)<5%, mild 5%–10%, moderate 15%–30%, severe 35%–70%.
UK Liver EQA participants steatosis assessment study

- **Aim:** to see if the subjective assessment of the steatotic proportion of biopsy parenchymal area (estimated fat proportionate area, eFPA) is influenced by:
  - The realisation that there is a prevalent tendency to systematically overestimate the steatotic proportion of biopsy parenchymal area;
  - And the provision of guideline images.
- **Two circulations** (different images in each circulation) of H&E images (x4 and x20 magnification) with range of steatosis.
- **Second circulation** accompanied by an article discussing the difference between eFPA and image analysis measurement of the area proportion of steatosis (mFPA) and guideline images demonstrating a range of mFPA.
  - 38 participants responded to circulation 1
  - 23 participants responded to circulation 2
- **Use of Kleiner et al. assessment system:**
  - Circulation 1: 18/38 (47.4%); Circulation 2: 11/23 (47.8%)
- **18/23 (78.3%)** participants who responded to circulation 2 used a printed out copy of the sample images
- **19 participants** responded to both circulations
EQA steatosis assessment study

1\textsuperscript{st} circulation eFPA>mFPA vs 2\textsuperscript{nd} circulation eFPA≈mFPA (esp. x4)
EQA steatosis assessment study

Kleiner users scores are almost the same as Kleiner non-users
EQA steatosis assessment study

2nd circulation: users of printouts of guideline images achieve almost the same FPA as image analysis.
EQA steatosis assessment study
19 participants responded to both circulations

Responders to both circulations:
1st circulation overestimates are corrected in 2nd circulation
# EQA steatosis assessment study

Circulation 2: ranking by correlation coefficient

<table>
<thead>
<tr>
<th>Ranking by correlation coefficient</th>
<th>x20 Objective Magnification</th>
<th></th>
<th>x4 Objective Magnification</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant EQA number</td>
<td>r</td>
<td></td>
<td>Participant EQA number</td>
<td>r</td>
</tr>
<tr>
<td>1</td>
<td>91</td>
<td>0.9804366</td>
<td>1</td>
<td>148</td>
</tr>
<tr>
<td>2</td>
<td>148</td>
<td>0.9791633</td>
<td>2</td>
<td>91</td>
</tr>
<tr>
<td>3</td>
<td>67</td>
<td>0.9754431</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>77</td>
<td>0.9689558</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>5</td>
<td>41</td>
<td>0.9666655</td>
<td>5</td>
<td>119</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>0.9639714</td>
<td>6</td>
<td>67</td>
</tr>
<tr>
<td>7</td>
<td>145</td>
<td>0.9571206</td>
<td>7</td>
<td>145</td>
</tr>
<tr>
<td>8</td>
<td>130</td>
<td>0.9566676</td>
<td>8</td>
<td>144</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.9525423</td>
<td>9</td>
<td>130</td>
</tr>
<tr>
<td>10</td>
<td>106</td>
<td>0.9490474</td>
<td>10</td>
<td>40</td>
</tr>
</tbody>
</table>
Liver biopsy assessment of steatosis
(in routine paraffin processed H&E sections)

Conclusion

• Mild vs moderate vs severe steatosis is arbitrary, subjective, idiosyncratic and undefined.
  • We need a better consensus regarding what constitutes the descriptive categories of mild vs moderate vs severe steatosis.
• Neither the mild/moderate/severe nor the ≤33%/33-66%/≥66% steatosis classification has yet been shown to have any reliable clinical relationship, and the cut-offs are of uncertain significance.
  • The use of 33% and 66% cutoffs for mild vs moderate vs severe steatosis is arbitrary.
  • The cutoffs should be based on clinical correlations and relevance.
• We are systematically overestimating the tissue sectional area of fat.
  • The tissue sectional area of fat scale is 0-50% (and not 0-100%).
• We will not know if liver biopsy assessment of steatosis has any practical clinical significance or not until we begin to make the assessment more accurately.
• But it is still important to assess liver biopsy steatosis properly so that the non-invasive methods for the diagnosis of fatty liver are calibrated correctly.
Grading fat according to % of hepatocytes that are steatotic

- There is no definition of “steatotic hepatocyte”.
  - It is a normal function of liver cells to metabolise and make fat.
  - What proportion of a hepatocyte's cytoplasm should be fatty for that hepatocyte to be considered “steatotic”, ie how much fat per hepatocyte is normal vs abnormal?

- % “steatotic” hepatocytes fat grading requires knowledge of:
  - How many hepatocytes are there altogether in each microscopic field?
  - How many of the hepatocytes in each microscopic field are steatotic?

- Enumerating hepatocytes properly is impossible at low magnification.

- “Numbers of hepatocytes” counting is practically impossible even at higher magnifications because:
  - Tissue sections are usually 4-5 µ.
  - Hepatocytes are >15 µ (up to 100µ in steatotic/ballooned hepatocytes).
    - In any given plane of section:
      - Many hepatocytes are cut tangentially.
      - Hepatocytes are in different planes of section.
      - Hepatocytes have overlapping and indistinct borders with adjacent cells.

- % of steatotic hepatocytes estimation from how much of the cross sectional area of each lobule is steatotic is not easy because fewer enlarged steatotic hepatocytes occupy the same area as more non-steatotic hepatocytes which are smaller.

- Enumerating % steatotic hepatocytes as an assessment in fatty livers is a myth that cannot be substantiated or sustained.

<table>
<thead>
<tr>
<th>mFPA</th>
<th>Median eFPA (range)</th>
<th>Steatosis grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10% (10–20%)</td>
<td>Mild 100%</td>
</tr>
<tr>
<td>15</td>
<td>25% (10–40%)</td>
<td>Mild 42%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate 58%</td>
</tr>
<tr>
<td>20</td>
<td>35% (20–40%)</td>
<td>Mild 42%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate 58%</td>
</tr>
<tr>
<td>25</td>
<td>60% (40–70%)</td>
<td>Moderate 75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severe 25%</td>
</tr>
<tr>
<td>30</td>
<td>60% (35–70%)</td>
<td>Moderate 66%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severe 33%</td>
</tr>
<tr>
<td>35</td>
<td>75% (60–80%)</td>
<td>Moderate 17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severe 83%</td>
</tr>
</tbody>
</table>

Subjective mild/mod/severe steatosis consensus is lost between 15-30% mFPA [eFPA (25 (10-40)->60% (35-70)%)]