Liver echogenicity: measurement or visual grading?

Tapio Vehmasa,*, Ari Kaukiainenb, Katariina Luomaab, Martina Lohmanb, Markku Nurminena, Helena Taskinenab

aThe Finnish Institute of Occupational Health (FIOH), Topeliuksenkatu 41 a A, FIN-00250 Helsinki, Finland
bHelsinki University Central Hospital, Peijas Hospital, Sairaalakatu 1, FIN-01400 Vantaa, Finland

Received 12 June 2003; revised 24 March 2004; accepted 24 March 2004

Abstract

Objective. Two methods to assess liver echogenicity were compared.

Methods. Liver/kidney echogenicity ratio was measured in 41 persons with the ultrasound software and visually graded by two radiologists and a radiographer. These echogenicity ratios and grades were related to risk factors for fatty liver and to liver enzyme levels.

Results. These determinants explained 55% of the radiologists’ mean grades, 14% of the radiographer’s and 31% of the measured echogenicity ratios.

Conclusion. Radiologists’ visual gradings correlated best with the indirect determinants of early liver pathology. Computerized measurements may be inferior to visual grading due to the lack of holistic tissue diagnostics.

Keywords: Echogenicity; Enzymes; Fat; Liver; Measurement; Observer; Ultrasound

1. Introduction

The increased echogenicity of liver, or ‘bright liver’, was recognized in the 1970s. A prevalence of 20% was reported from Italy [1]. Body mass index (BMI), age, serum cholesterol, triglycerides and apo B levels [1], and hypertransaminasemia [2] are all independent predictors of a bright liver. In recent studies non-alcoholic fatty liver has been found to be associated with insulin resistance [3] and the metabolic syndrome [4]. The increase of liver echogenicity is caused mainly by fatty degeneration, and fibrosis is a minor contributor [5,6]. The sensitivity of ultrasound (US) to detect liver fatty degeneration was 89% [7] and the method has been considered superior to computed tomography [8].

Liver echogenicity is usually estimated by comparing it to that of the right kidney cortex. Subjective estimation is observer dependent and may be inaccurate. Densitometry has been regarded indispensable for the accurate density assessment due to the suboptimal potential of human eye to detect small differences in optical density [9]. In this study we compared the quantitative estimation of liver echogenicity to visual assessment. The determinants known to be associated with early liver pathology (liver enzymes) or liver fatty degeneration (BMI, serum triglyceride level) were regarded as indirect determinants of liver changes and used as a gold standard in this study.

2. Subjects and methods

2.1. Subjects

The study group consisted of 22 workers occupationally exposed to solvents (3 women and 19 men, mean age 52 years, range 29–66 years) and 19 non-exposed persons (4 women and 15 men, mean age 45 years, range 30–57 years), who were employees of the institute (FIOH). The exposed persons were initially referred to medical examinations because of suspected solvent-related chronic toxic encephalopathy or polyneuropathy. Exclusion criteria included systemic diseases or medication with known...
hepatic effects, positive hepatitis serology and current pregnancy. BMI varied between 19 and 36 (mean 26) kg/m². This study was conducted with the approval from the local ethics committee. All subjects participated voluntarily and provided informed consent.

2.2. Liver sonography and laboratory tests

The SonoAce 8800 MT ultrasound scanner (Medison, Soul, South Korea; http://www.trimedic.com/ult/8800.htm) and the convex (3.5–5 MHz) probe were used. The histogram feature of this scanner is capable of measuring 256 shades in the grey scale. The time-gain compensation was set to adjust the tissue echogenicity as constant as possible regardless of depth. Two foci were used. A representative longitudinal US slice showing the liver parenchyma and the neighboring right kidney was selected for the measurement of echogenicity by a radiologist (TV) and printed with Mitsubishi P 91E thermic printer (Mitsubishi, Kyoto, Japan) for blinded visual grading. The echogenicity of the liver parenchyma was measured in three different regions of interest (ROI) close to the adjacent right kidney in the same slice and depth, avoiding vessels. Similar measurements were performed in the kidney cortex (Fig. 1). Mean values were accepted. The echogenicity ratio (mean liver echogenicity/mean kidney echogenicity) was then calculated.

The degree of liver echogenicity was also visually assessed by two experienced radiologists (KL, ML) and by a radiographer with little experience in sonography. They compared the echogenicity of the liver to that of the adjacent kidney cortex in the printed images by using reference images [10] (scale: normal = 1, mild abnormality = 2, severe abnormality = 3). The observers used decimals if they could not match the particular US image with any of the reference images. Both radiologists’ mean grade was also calculated for each subject.

Blood samples were obtained in the morning after fasting for 12 h, before liver sonography was performed. They included serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase (GGT), carbohydrate-deficient transferrin (CDT) and triglycerides. Standard laboratory methods were used.

2.3. Statistical methods

The agreement between the three repeated echogenicity measurements and, on the other hand, the agreement between the readers’ grades was assessed by using the intraclass correlation coefficient (ICC). Measured (liver/kidney) echogenicity ratios and readers’ grades were compared with Pearson’s correlation.

The relation between the indirect determinants of early liver pathology and liver echogenicity was studied using linear regression modeling. To comply better with the assumption of the normality of the dependent echogenicity variables (e.g. the symmetry of the distribution) we subjected them to square root transformation. The determinants were entered into the model as quadratic variables because of the non-linearity of some of the relations. From the regression analysis results we calculated the coefficient of determination, that is, the square of the product-moment correlation between two variables, $r^2$. This quantity expresses the proportion of the variance of one variable given by the other [11]. We related both

Fig. 1. Echogenicity measurement. A representative area from liver close to the right kidney (1M:31) and the kidney cortex at the same depth (2M:28) show echogenicities.
the radiologists’ echogenicity grades separately, their mean values and the measured echogenicity ratios to the indirect determinants of early liver pathology by using linear multiple regression modeling. In the model, the grade and the ratio had a multiple correlation, \( R \); \( R^2 \) indicates the fraction of the total variation in the liver echogenicity, which is accounted for by the regression model. S-PLUS 2000 system was applied for data analysis (MathSoft, Inc., Seattle, WA, USA, 1999).

3. Results

The ICC between the three repeated echogenicity measurements was 0.93 for the liver and 0.94 for the right kidney. The ICC of the visual grades was 0.65 between radiologists 1 and 2, 0.45 between radiologist 1 and the radiographer, and 0.37 between radiologist 2 and the radiographer.

Correlations between the measured (liver/kidney) echogenicity ratio and the readers’ grades were: 0.55 \( (p < 0.001) \) for radiologist 1, 0.28 \( (p = 0.07) \) for radiologist 2, and 0.02 \( (p = 0.9) \) for the radiographer. There were slight differences between the readers’ mean grades of echogenicity (radiologist 1: 1.92, radiologist 2: 1.79, radiographer: 1.69). The average grade of the radiologists was plotted as a function of the measured echogenicity ratio (Fig. 2).

Coefficients of determination between echogenicity (estimated by the readers or measured with US) and the indirect determinants of early liver pathology (study subjects’ BMI, serum triglyceride and liver enzyme levels) are given in Table 1. The measured echogenicity was poorly explained by BMI, triglycerides, and CDT, while the readers’ echogenicity grade was better explained by ALT, AST, and best by BMI. The radiologists’ grades correlated better with the indirect determinants of early liver pathology than the radiographer’s grades. The mean grades of the two radiologists correlated better with these determinants than individual radiologists’ grades.

4. Discussion

The radiologists’ echogenicity grades showed better correlation with the indirect determinants of early liver pathology than did the echo ratio measurements or the grades of the radiographer. This fact was especially pronounced in relation to the BMI.

Liver biopsies were not clinically indicated and thus could not be obtained for ethical reasons. Due to the lacking histological gold standard we compared our echogenicity findings with the combination of variables previously known to be associated with fatty liver or non-specific liver parenchymal damage. We did not consider histology mandatory, because our aim was to compare the visually based and computerized US methods with each other and not to study associations between liver histology and US.

The measured echogenicity deep in tissue reflects both the intrinsic echogenicity of the area itself and the acoustic properties of all tissues between the probe and this area. Especially in obese patients, the heavy contribution of

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect determinants of early liver pathology in relation to its echogenicity</td>
</tr>
<tr>
<td>Echogenicity grade/measurement</td>
</tr>
<tr>
<td>Radiologist 1</td>
</tr>
<tr>
<td>Radiologist 2</td>
</tr>
<tr>
<td>Mean of radiologists 1 and 2</td>
</tr>
<tr>
<td>Radiographer</td>
</tr>
<tr>
<td>Mean of all readers</td>
</tr>
<tr>
<td>US measurement (liver/kidney ratio)</td>
</tr>
</tbody>
</table>

Coefficients of determination \( (R^2) \) between liver echogenicity (estimated by the readers or by measured echo ratio) and body mass index (BMI) and laboratory tests (trigly, serum triglycerides; AST, aspartate amino transferase; ALT, alanine amino transferase; GGT, gamma glutamyl transferase; CDT, carbohydrate-deficient transferrin).
superficial tissues may cause unreliability of measurement. Our good ICCs between repeated echogenicity measurements suggest that this method may theoretically be the best to quantify the shades of grey according to what Smith-Levitin et al. [9] have found. However, this technical ability may not be sufficient for reliable grading of liver echogenicity. Computerized measurements are restricted to ROI areas, which are usually set small to exclude disturbing structures such as vessels, bile ducts or acoustic shadows from other tissues. On the contrary, the human eye may register the whole liver parenchyma and kidney cortex without artificial ‘sampling’ of tissue, in a more holistic manner. Medically trained observers may also be able to exclude the effect of artifacts. In addition to echogenicity, the reference images focused also to the visibility of vascular markings and the diaphragm. Our observers could take into account the effect of these features and of liver size in their evaluation.

Our radiologists had better results than the radiographer did. This observation contrasts the finding of Smith-Levitin et al. [9], who found no difference in echogenicity estimation due to medical qualification. Their study was designated to compare mere tissue echo density, however.

As a conclusion, computerized measurements of liver echogenicity suffer from the lack of an integrated diagnostic approach, which can be reached only by medically trained observers. US echogenicity measurements should therefore, be dealt with care, even if they are echo ratios. Experienced observers’ grading is currently the preferred method for ultrasonic evaluation of liver echo texture.

5. Summary

We compared the visually based and computerized measurement method for assessing the liver echogenicity and its minimal changes. The echogenicity of the liver and adjacent kidney cortex was measured in 41 persons (aged 29–66 years) with the US software to determine the liver/kidney echo ratio. Also, two radiologists and a radiographer graded liver echogenicity visually by using reference images. The echo ratios and visual grades were related to risk factors for fatty liver (BMI and the level of serum triglycerides) and to liver enzymes, the combination of which was used as an indirect gold standard of early liver pathology. These determinants explained 48% of both radiologists’ visual gradings, 55% of the radiologists’ mean gradings, 14% of the radiographer’s gradings, and 31% of the measured echogenicity ratios. The radiologists’ visual grades were thus best indicators of early liver pathology. Local echo artifacts especially in obese individuals and the lack of holistic tissue diagnostics may impede computerized echogenicity measurements. An experienced radiologist pays attention to vascular architecture and the effect of liver size, diaphragmatic visibility and artifacts in addition to the general echogenicity of liver. Experienced observers’ grading is currently the preferred method for the diagnostic US evaluation of early liver abnormality.

Acknowledgements

We wish to thank radiographer Elise Koskenseppä for estimating liver echogenicity and Ms. Terttu Kaustia, MA, for language revision.

References

Tapio Vehmas completed his MD from Helsinki University in 1984. He received his specialist degree in radiology in 1991 and published his doctoral thesis work in the same year. He has worked in the Helsinki University in 1991–1993 as a lecturer (assistant professor) teaching radiology and in the Helsinki University Central hospital as a senior radiologist during 1991–1998. He was appointed to a senior lecturer (docent) of radiology in Helsinki University in 1997. Since 1999 he has worked as the chief radiologist at the Finnish Institute of Occupational Health.

Ari Kaukiainen completed his MD from Helsinki University in 1989 and received his specialist degrees in occupational health in 1998 and occupational medicine in 2000. Since 1998 he has worked at the Finnish Institute of Occupational Health, and since 2001 in charge of a work ability assessment unit at the Department of Occupational Medicine. He is currently preparing his doctoral thesis about health effects of organic solvents and paint compounds.

Martina Lohman completed her MD from the University of Helsinki in 1986. She made a two-year residency in the field of surgery and orthopedics before shifting to radiology. In 1995 she completed her specialization in radiology, and published her doctoral thesis in 2001. From 1996 to 2002 she worked as a senior radiologist at the Helsinki University Central Hospital. During 2002–2003 she was appointed as a Clinical Assistant Professor at the University of Michigan, USA. Her publications are mainly in the field of musculoskeletal MRI.

Katriina Luoma completed her MD from Tampere University in 1979 and received her specialist degree in radiology in 1986 and in neuroradiology in 2002. During 1992–1993 she worked at the Finnish Institute of Occupational Health. She published her thesis work in 2000. She has worked as a senior radiologist during 1991–1996 and since 1997 as the deputy head of department at the Department of radiology, Peijas Hospital, Helsinki University Central Hospital. Her scientific interest is mainly focused on MRI and epidemiology of musculoskeletal disease.

Markku Nurminen works as a senior research scientist in the Department of Epidemiology and Biostatistics, Finnish Institute of Occupational Health. He holds the position of senior lecturer (docent) in biometry at the Department of Public Health, University of Helsinki, being a specialist in biostatistics and epidemiological methods. His research has focused on the field of occupational and public health for 30 years, and recently on estimating working life expectancies. He is a (co)author of about 150 publications.

Helena Taskinen is the director of the Department of Occupational Medicine at the Finnish Institute of Occupational Health, and Professor of Occupational Health at the School of Public Health at the University of Tampere, Finland. Her research experience covers occupational reproduction epidemiology, occupational diseases (solvent induced diseases, musculo-skeletal diseases of hand and arm from repetitive work, respiratory diseases of workers at water damaged buildings), work conditions and well being of pupils and personnel at schools, training and education of occupational health personnel and work conditions, as well as health and occupational services of contingent workers.