Clinical applications of Dixon chemical shift MR imaging: Morbus Gaucher, Morbus Hansen
Maas, M.

Citation for published version (APA):
Ridderkerk: M. Maas

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Dixon quantitative chemical shift MRI for bone marrow evaluation in the lumbar spine: A reproducibility study in healthy volunteers

Mario Maas, Erik M. Akkerman, Henk W. Venema, Jaap Stoker, Gerard J. den Heeten

Department of Diagnostic Radiology, Academic Medical Center, Amsterdam, the Netherlands

*J Comput Assist Tomogr* 2001; 25(5):691-697
ABSTRACT

Purpose:
The purpose of this work was to explore the reproducibility of fat-fraction measurements using Dixon quantitative chemical shift imaging (QCSI) in the lumbar spine (L3, L4, and L5) of healthy volunteers.

Method:
Sixteen healthy volunteers were examined at 1.5 T two times to obtain a repeated measurement in the same slice and a third time in three parallel slices. Single slice, two point Dixon spin echo (TR/TE 2,500/22.3) sequences were used, from which fat-fraction images were calculated. The fat-fraction results are presented as averages over regions of interest, which were derived from the contours of the vertebrae. Reproducibility measures related to repeated measurements on different days, slice position, and contour drawing were calculated.

Results:
The mean fat-fraction was 0.37 (SD 0.08). The SD due to repeated measurement was small ($\sigma_R = 0.013$ to 0.032), almost all of which can be explained by slice-(re)-positioning errors.

Conclusion:
When used to evaluate the same person longitudinally in time, Dixon QCSI fat-fraction measurement has an excellent reproducibility. It is a powerful non-invasive tool in the evaluation of bone marrow composition.
Reproducibility QCSI in healthy volunteers

INTRODUCTION

The three major constituents of bone marrow in healthy individuals are mineralized osseous matrix, red cellular hematopoietic marrow, and yellow fatty marrow [1]. Red and yellow marrow are interconvertible [2], and normally they are in a dynamic balance that is age and gender dependent [3,4]. Disorders in this balance may occur following a number of pathophysiologic mechanisms: reconversion of yellow into red marrow (e.g., anemia, myeloproliferative diseases), marrow infiltration or replacement (e.g., neoplasm, infection, myelofibrosis, Gaucher disease), and myeloid depletion of red marrow (aplastic anemia, radiation therapy, chemotherapy) [1,5,6].

MRI is sensitive to both the water and the fatty components in bone marrow. Conventional qualitative MR images represent the sum of the signals from water and fat. In case of a disturbed balance of red and yellow marrow, the relative contribution of the fat signal to the total signal, known as "fat-fraction", may be a useful diagnostic quantitative parameter [7]. In Dixon quantitative chemical shift imaging (QCSI), based on the phase contrast technique described by Dixon [8], the MR signal is separated into the individual contributions of fat and water [9], from which the Dixon QCSI fat-fraction \( F_f \) is calculated. \( F_f \) has been demonstrated to be an excellent discriminator between healthy subjects, subjects with aplastic anemia, and subjects with leukemia [9]; furthermore, it has been shown to be useful in monitoring the extent of disease, disease progression, and response to therapy in acute leukemia [10,11] and Gaucher disease [12,13].

Although Dixon’s principal idea was published back in 1984 [8], interest in the method has never ceased. Among the more recent publications are application of gradient echo instead of spin echo in the imaging of bone marrow [14], quantitative assessment of water fraction and T2 in bone marrow [15], and the use of Dixon imaging as a fat-suppression method in hands and feet [16] and in children [17].

Relatively little, however, has been published about the reproducibility of Dixon QCSI in vivo. Buxton et al. [18] described the influence of some technical parameters at 0.6 T. Gückel et al. [19] conducted a serial measurement of gluteal muscle at 1.5 T. Rosen et al. [9] elaborated on the overall reproducibility of serial measurements as measured in three healthy volunteers at 0.6 T.
We conducted the current study to explore the reproducibility issue for the measurement of $F_j$ in the vertebrae L3, L4, and L5 at 1.5 T in a group of healthy volunteers in a clinical setting, that is, using exactly the same procedures as used in patients. Our main research question was: What is the reproducibility of $F_j$ when repeating the measurement on a different day? We also investigated the influence on the reproducibility of two operator-dependent aspects of the method: (re-)positioning of the measurement slice and determination of contours from which the regions of interest (ROIS) were calculated. Furthermore, we evaluated the available Dixon QCSI data in literature, compared these with our results, and evaluated the differences.

**MATERIAL & METHODS**

**Subjects**
For the measurements, we used a group of 16 healthy volunteers (8 men: mean age 39 years, range 24-60 years; 8 women: mean age 38 years, range 26-55 years). There was no history of bone marrow pathology in this group. Approval of the local medical ethics committee was obtained, and all volunteers gave their written informed consent.

**Experimental Procedures**
Every volunteer was examined on 3 different days. In two of the examinations, a fat-fraction measurement was performed in the same single slice, referred to as measurements a and b; together they are called the repeated measurement. The average interval of the two examinations was 43 days (range 7-155 days, median 28 days). During a third examination, fat-fraction measurements were performed in three parallel slices, referred to as measurements c1, c2, and c3; together they are called the threefold measurement. All procedures took place at our institution, and the MR examinations were performed on a 1.5 T Magnetom 63SP/4000 later upgraded to a 1.5 T Magnetom Vision (Siemens, Erlangen, Germany). Volunteers were examined either entirely on the Magnetom SP ($n = 9$) or entirely on the Magnetom Vision ($n = 7$). A Dixon fat-fraction measurement sequence consisted of two spin echo sequences. The first one was a conventional spin echo sequence, in which the signals from fat and water are in-phase. The second sequence was identical to the first one, except that the $180^\circ$ inverting radio-frequency
Pulse was shifted by a time \( \tau \), such that in the center of the readout gradient \( (k_x = 0) \), the signals of water and fat had opposed-phases \([8,20,16]\). The value of \( \tau \) depends on the field strength and on the frequency difference between fat \((-\text{CH}_2-)\) and water \((\text{H}_2\text{O})\), which is approximately 3.4 ppm [21]. At 1.5 T, this corresponds with a value of \( \tau = 1.15 \) ms.

In-phase and opposed-phase proton density-weighted Dixon spin echo sequences were performed with the following parameters: TR 2,500 ms, TE 22.3 ms, slice thickness 4 mm, matrix 256 x 256, NEX 1, FOV 350 x 350 mm, resulting in a pixel size of 1.37 x 1.37 mm. The sequences had a bandwidth of 130 Hz/pixel. The acquisition time for one spin echo sequence was 10 min 40 s and for a Dixon measurement sequence (in-phase and opposed-phase) 21 min 20 s. Only single slice acquisition was performed, with no presaturation slabs, to avoid errors due to cross-talk and off-resonance effects [9]. Since presaturation slabs could not be used, both the RF coil and the direction of the measurement slice had to be selected with care to avoid motion artifacts.

![Figure 1](image-url)

**Figure 1.** Mid-sagittal localizer images of volunteer v03. Measurement slices a and b (repeated measurement) and c1, c2, and c3 (threelfold measurement) are indicated on the images, and each slice is perpendicular to the localizer images. Measurement slices should pass through the junction of the anterior three-fourths with the posterior one-fourth of vertebral body L4. Slight angulation from coronal toward transversal is allowed to also make the slices pass, as best one can, through the posterior parts of L3 and L5. To improve the reproducibility of the slice positioning, previous localizer images always were consulted for subsequent slice positioning in the same volunteer.
We used (para)-coronal slices, which were positioned on a mid-sagittal localizer image, passing through the middle of the posterior parts of L3, L4, and L5, as illustrated in Fig. 1. The distance between the slices of the threefold measurement was 3 or 4 mm, depending on the size of the vertebrae.

On the Magnetom SP, the body Helmholtz coil was used; on the Magnetom Vision, we used the small flex coil, flattened out and placed underneath the body along the spine. In this way, the intensity distribution in the images as well as the signal-to-noise ratio of both coil-scanner combinations were similar. We did not use a phased array coil because our sequences could not provide useful phase images for phased array coils.

**Postprocessing**

Postprocessing was performed on a Sun Sparc 20-51 workstation (Sun Microsystems, Mountain View, CA, U.S.A.). The postprocessing steps necessary to obtain a fat-fraction image from the Dixon acquisitions are illustrated in Fig. 2. In the in-phase image, the signals from water and fat are added together (Fig. 2a). The information about the difference of the signals is contained in the opposed-phase image (Fig. 2b) and the phase angle of the acquired signals (Fig. 2c). Using an algorithm earlier described [16,22] a pixel-to-pixel, region-growing version of a method by Brix et al. [23] we obtained water-only (Fig. 2e) and fat-only (Fig. 2f) images. Due to the chemical shift displacement effect, there is a fat-water shift of 1.7 pixels in the readout (head-feet) direction. We corrected for this by shifting back the water-only image relative to the fat-only image by 2 pixels. This correction removed almost all of the displacement artifacts in the fat-fraction image (Fig. 2g), which would have been present otherwise. The fat-fraction image was calculated using the following formula:

\[
F_f = \frac{M_f}{M_f + M_w}
\]

- 26 -
Figure 2. Postprocessing of a Dixon sequence. Shown are the in-phase magnitude image (a), opposed-phase magnitude image (b), and image of the phase difference between in-phase and opposed-phase sequences (c). From (c), the sign information (d) is obtained: In the lighter parts, water signal > fat signal; in the darker parts water signal < fat signal. Combining (a), (b) and (d), the water image (e) and the fat image (f) are obtained. Parts (e) and (f) have been shifted two pixels in the vertical direction with respect to each other, so as to correct for the chemical shift displacement. From (e) and (f), we calculate the fat-fraction image (g). Also shown is the scale relating gray values in (g) to fat-fraction numbers (h). Note: Postprocessing is done with rectangles of 64 x 128 pixels, as displayed here, containing L3, L4, and L5, which have been cut out of the original 256 x 256 images.
Figure 3. Determination of regions of interest. In the in-phase magnitude image (a), the vertebrae are clearly distinguishable from their surroundings. Interactively, with the help of a mouse, the contours of the vertebrae L3, L4, and L5 are drawn (b). By applying three erosion operations to the interior region of the contours, the regions of interest (c) are obtained.

in which $F_f$ denotes the pixel values in the fat-fraction image, $M_i$ the pixel values in the fat-only image, and $M_w$ the pixel values in the water-only image. In this way, for every pixel, an $F_f$ value was obtained, between 0 (no fat) and 1 (only fat, no water).

To obtain one fat-fraction value for each vertebra, we averaged the pixel values in an ROI. Figure 3 shows how these ROIS were obtained. In the in-phase image, the contours of the vertebrae L3, L4, and L5 were manually drawn, using a mouse. ROIs were calculated by applying three erosions to the region enclosed by the contours (Fig. 3c). "Erosion" is an image-processing procedure in which the edge pixels of a region are removed [24]. In this way, the outer parts of the vertebrae were excluded, which was done to avoid partial volume effects and sequelae of secondary bony change in degenerative disc disease [9,25]. ROIs with <100 pixels were considered non-representative and were discarded. To assess the influence of contour drawing on $F_f$, a second observer independently drew contours in the images of measurement a.
Reproducibility QCSI in healthy volunteers

**Figure 4.** Data plot of the threefold measurement. x axis: mean $F_j$; y axis: differences of measurement data with the individual means ($\Delta F_j$). Squares: measurement a; triangles: measurement b; dashed lines at $\pm \sigma_R$; SD of all points; dotted lines at $\pm 2\sigma_R$: 95% confidence intervals.

**Figure 5.** Data plot of the threefold measurement. x axis: mean $F_j$; y axis: differences of measurement data with the individual means ($\Delta F_j$). Squares: measurement c1; triangles: measurement c2; diamonds: measurement c3; dashed lines at $\pm \sigma_R$; dotted lines at $\pm 2\sigma_R$: 95% confidence interval.

In the analysis of the repeated measurement, we assumed that a specific measured $F_j$ value was drawn from a normal distribution $N(\mu, \sigma_R)$, where the average value $\mu$ is the true $F_j$ value for that volunteer and the SD $\sigma_R$ is the error associated with repeating the measurement on a different day. $\sigma_R$ was taken as our measure for the reproducibility and was derived by
plotting the data as described in Results (Fig. 4) similar to that previously described [26,27]. In the same way, we derived $\sigma_r$, the SD associated with the slice-repositioning error from the threefold measurement (Fig. 5), and $\sigma_c$, the SD associated with interoperator contour drawing variability, from the results of the two contours. All calculations were performed for the three vertebrae separately. The significance of differences was tested with a two sided paired $t$ test. Results were considered significant if $p < 0.05$.

**Results**

All volunteers underwent the procedures without problems. Every examination was adequate. In one volunteer, there was a significant scoliosis of the spine, which made it somewhat difficult to obtain a mid-sagittal localizer slice for accurate positioning of the measurement plane. Three volunteers had a relatively large lordosis, causing the ROI of L5 in the most anterior slice of the threefold measurement to be smaller than 100 pixels. Therefore, these data were discarded. The remaining ROIs had an average (range) of 254 (153-343) pixels in the repeated measurement and 242 (128-341) pixels in the threefold measurement. We observed no consistent differences between the results from the two coil-scanner combinations.

In Fig. 4, we show the results of the repeated measurement: on the $x$ axis, the individual mean $F_r$ values of measurements a and b are plotted (as an estimate of the "true" individual value $\mu_i$) and on the $y$ axis, the differences of the $F_r$ values with these means ($\Delta F_r$). The SD in the $y$ direction is $\sigma_r = 0.032$ in L3, 0.020 in L4, and 0.013 in L5. We found no significant differences between the $F_r$ values of measurements a and b. The group averaged $F_r$ (SD) values were 0.37 (0.08) in L3 and L4, and 0.38 (0.08) in L5.

In the same manner, the results of the threefold measurement are shown in Fig 5. For the SDs, we find $\sigma_r = 0.028$ in L3, 0.023 in L4 and 0.015 in L5. We found no significant differences between the $F_r$ values of measurements c1, c2, and c3.

There were no significant differences between the results from the contours from two operators, as performed in measurement a. The associated reproducibility measure, $\sigma_c$, was 0.003 in L3, 0.002 in L4, and 0.002 in L5.


**DISCUSSION**

Dixon QCSI is an established method for the characterization of various bone marrow disorders [18,19,25]. It has been reported as a powerful non-invasive tool for the quantitative analysis of bone marrow invasion [9-12,28], and it is stated to be the most sensitive technique to evaluate bone marrow response to the expensive enzyme supplementation therapy in Gaucher disease [29]. In view of the small number of evaluation studies of this method, we performed the present measurements to assess a number of reproducibility aspects in a group of healthy volunteers. We also compare our results with the available literature data.

**Reproducibility**

The $\sigma_r$, the reproducibility of $F_r$, when the measurement is repeated on a different day, comprises many sources of error, among which are physiological variabilities, machine instabilities, slice-positioning errors, and contour drawing (ROI) variations. The two latter variables are operator dependent and therefore were analyzed in detail.

The contribution of contour drawing to $\sigma_r$ is very small. Since the $\sigma$ values are SDs, they add quadratically, and therefore the contribution of $\sigma_c$ to $\sigma_r$ is negligible. Apparently, the procedure for obtaining ROIs is very reproducible. Slice positioning was performed on an (interpolated) localizer image with an anteroposterior resolution of about 4 mm, and we do not expect the positioning accuracy in the repeated measurement to be much smaller. In the threefold measurement, we deliberately varied the slice position by 3-4 mm, which is comparable with the estimated positioning variance. From the fact that $\sigma_t$ and $\sigma_r$ are of the same magnitude, we conclude that the contribution of the error due to slice (re)-positioning to $\sigma_r$ is considerable. Paying attention to accurately repositioning the measurement slice is of crucial importance to the reproducibility; a localizer image with a higher resolution is likely to give improvement.

We did not analyze in detail the physiological variability, instabilities in the equipment, and other causes for error and variations that may contribute to $\sigma_r$. However, $\sigma_r$ is the upper limit; whatever the effect of the physiological variability on this measurement, it could never have been larger than $\sigma_r$. Since we have seen already, that slice positioning can
explain the larger part of $\sigma_R$, these other influences can be of only minor importance.

An interesting observation from the $\sigma$ values and Figs. 4 and 5 is that both in the repeated measurement ($\sigma_R$) as well as in the threefold measurement ($\sigma_T$), L3 has the worst reproducibility. When reviewing the images, there appeared to be more artifacts in L3, probably owing to respiration or blood flow in vessels, causing artifacts in the phase direction (left-right).

Comparison with literature data on healthy volunteers

We found only one study in the literature that included repeated measurements with Dixon QCSI in the lumbar vertebrae: Rosen et al. [9] report fat-fraction measurements at 0.6 T in three volunteers, repeated two or three times, giving a total of only seven measurements. Recalculation of their data using our analysis method yields $\sigma_R = 0.023$, which is comparable to the results of our study: $\sigma_R = 0.013$ to 0.032.

Dixon QCSI fat-fractions of healthy volunteers without repetition of measurements are reported more frequently: Table 1 lists our data together with data from the literature with $F_r$ values ranging from 0.29 to 0.48. To understand the apparent differences between the various studies, the methodological and physiological factors that might affect the measured $F_r$ values should be considered.

The most important methodological factors that vary in these studies are TR, the field strength, and whether relaxation correction is applied. Water T1 is considerably longer than fat T1, so T1-weighting causes a stronger reduction of the water signal than of the fat signal, and therefore a shorter TR will lead to a higher measured $F_r$. This might explain the high $F_r$ values found previously [19,25,28], for these authors used the shortest TR (1,200 ms); their mutual differences are small compared with the reported SDs. At a lower field strength, T1 relaxation is faster, and hence the measured $F_r$ will be lower. By explicitly applying relaxation correction, which implies performing T1-, T2- and proton density-weighted Dixon sequences, one is insensitive to TR, TE and $B_0$. 

- 32 -
Reproducibility QCSI in healthy volunteers

**TABLE 1.** Average fat-fraction values of healthy volunteers, measurement parameters, age and gender

<table>
<thead>
<tr>
<th>Ref.</th>
<th>$B_0$ (T)</th>
<th>RC*</th>
<th>TR/TE (ms)</th>
<th>n</th>
<th>Gender</th>
<th>Age (yrs)</th>
<th>Vertebral</th>
<th>$F_f$</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>This study</td>
<td>1.5</td>
<td>No</td>
<td>2,500/22.3</td>
<td>16</td>
<td>8M, 8F</td>
<td>24-60</td>
<td>L3-5</td>
<td>0.37</td>
<td>0.09</td>
</tr>
<tr>
<td>21</td>
<td>1.5</td>
<td>No</td>
<td>2,000/22</td>
<td>3</td>
<td>1M, 2F</td>
<td>18-21</td>
<td>L4-5</td>
<td>0.33</td>
<td>0.04</td>
</tr>
<tr>
<td>19</td>
<td>1.5</td>
<td>No</td>
<td>1,200/22</td>
<td>9</td>
<td>?</td>
<td>19-49</td>
<td>L4-5</td>
<td>0.48</td>
<td>0.10</td>
</tr>
<tr>
<td>25</td>
<td>1.5</td>
<td>No</td>
<td>1,200/22</td>
<td>9</td>
<td>?</td>
<td>12-49</td>
<td>L3-5</td>
<td>0.41</td>
<td>0.13</td>
</tr>
<tr>
<td>28</td>
<td>1.5</td>
<td>No</td>
<td>1,200/22</td>
<td>17</td>
<td>?</td>
<td>18-52</td>
<td>L3-5</td>
<td>0.44</td>
<td>?</td>
</tr>
<tr>
<td>12 (9)</td>
<td>0.6</td>
<td>Yes</td>
<td></td>
<td>6</td>
<td>?</td>
<td>21-47</td>
<td>L2-5</td>
<td>0.29</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*RC = relaxation correction

In fact, this corresponds to a proton density-weighted sequence with TR $\rightarrow \infty$ and TE $\rightarrow 0$, and the resulting $F_f$ will be lower than without relaxation correction. This explains the low $F_f$ value at 0.6 T reported in Table 1. Disadvantages of relaxation correction, however, are increasing imaging time and performing additional calculations, which often cause problems [9,10].

Important physiological factors known to influence $F_f$ are age and gender. Ishijima et al. [4] reported fat-fractions in females slowly increasing with age from 5 to 44 years and rapidly increasing in women older than 45 years; for males, a rapid increase of fat-fraction in the age group of 5-25 years with virtually no increase above 25 years was reported. The low $F_f$ values reported previously [18] might be explained by their very young volunteers. The relations we find (Fig. 6) are in accordance with those reported by Ishijima et al. [4], with no correlation of $F_f$ with age in the male group (ages 24-60 years) and a strong correlation in the female group (ages 26-55 years).
Clinical Implications
The $\sigma_r$ (0.013-0.032) we found is quite small compared with the influence of pathology on $F_r$ as found in the literature. The reported changes in $F_r$ values due to various hematologic pathologies as well as Gaucher disease range from 0.19 to 0.35 [10,12,25,28]. Furthermore, $\sigma_r$ is three to four times as small as the SD of $F_r$ in our population: 0.08.

Our results support the use of this technique as a non-invasive parameter in various hematologic or other bone marrow pathologies as a parameter of response to therapy or as a detector of relapse of disease. The use of fast or turbo spin echo may shorten the imaging time [30]. This has been explored, however, only for fat-suppression purposes; there is no experience yet with its use in QCSI.

One must remain careful in comparing results between studies and centers. It seems necessary to set up a database of normal population with age/gender values specific for the MR parameters used and the properties of the sequences used in every institution that wishes to implement this technology. Furthermore, every patient who will be studied repeatedly in time should be scanned with all technical parameters kept constant.
Reproducibility QCSI in healthy volunteers

CONCLUSIONS
Dixon QCSI fat-fraction measurement, when repeated on different days, has a very good reproducibility: \( \sigma_R = 0.013 \) (in L5) to 0.032 (in L3). \( \sigma_R \) is quite small compared with the changes due to pathology as found in literature. Differences in the fat-fraction between various vertebrae in the same volunteer are very small.

The contribution of contour drawing variations to the reproducibility is negligible. The slice-repositioning error, however, is of some importance. The use of a high resolution localizer image seems to be advisable.

Taking into account the influence of age, gender, and technical factors on the fat-fraction, our fat-fractions are in accordance with the literature, as are the age and gender dependencies we found. However, one must be careful when comparing fat-fraction values between centers.

The most valuable application of Dixon QCSI fat-fraction measurements in the lumbar vertebrae is in following the same person longitudinally in time, keeping all technical parameters constant. In that case, one profits most from the good reproducibility; pathologic variations as well as the effects of treatment are most likely to show. The inclusion of this method in research protocols concerning red bone marrow involvement or marrow characterization is therefore highly recommended. Its potential in the clinical evaluation of bone marrow pathology is a subject of ongoing study.

ACKNOWLEDGEMENT
The authors thank Carla E.M. Hollak M.D., Ph.D., and Johannes M.F.G. Aerts, Ph.D., for their support in executing this study.

REFERENCES


