



Published in final edited form as:

*J Clin Densitom.* 2019 ; 22(3): 329–337. doi:10.1016/j.jocd.2018.07.001.

## Vertebral Bone Density Measurements by DXA are Influenced by Hepatic Iron Overload in Patients with Hemoglobinopathies

Haven M. Allard<sup>1,\*</sup>, Lisa Calvelli<sup>1,2</sup>, Marcela G. Weyhmler<sup>1,2</sup>, Ginny Gildengorin<sup>1</sup>, and Ellen B. Fung<sup>1,2</sup>

<sup>1</sup>Children's Hospital Oakland Research Institute, Oakland, CA, USA.

<sup>2</sup>Department of Hematology, UCSF Benioff Children's Hospital, Oakland, CA, USA

### Abstract

Interference from metal hardware (piercings; buttons on clothing; and ingested material, e.g. barium) is well documented in bone health assessments by dual-energy X-ray absorptiometry (DXA). It is unknown if iron in hepatic tissue of highly iron-loaded patients could be mistakenly assessed by DXA as bone, and if this would lead to increased areal bone mineral density (aBMD) lumbar spine *Z-scores* derived by DXA. Our hypothesis is that iron in the liver of heavily loaded patients will artificially raise aBMD in the spine, and thereby lead to an error in the DXA scan. This study consisted of a retrospective chart review and re-analysis of DXA scans from patients with sickle cell disease and thalassemia combined with prospective DXA and liver iron concentration (LIC) measurements from healthy controls. Patients who previously had both a DXA and LIC measurement were compared with controls. aBMD of individual vertebrae were analyzed and grouped by those that may be covered by the liver (L1 or L1/2) with those typically not (L3/4). Subjects were grouped by diagnosis and LIC severity. Phantoms were created to mimic the geometry of iron loaded liver tissue, and analyzed by DXA. A significant effect was observed in the difference of BMD Z-score of L1 and L 3/4 when patients with LIC < 1000 were compared to those with >5000  $\mu\text{g Fe/g}$  wet tissue ( $p = 0.043$ ). A significant relationship was also observed in the difference in aBMD Z-score of L1 and 3/4 when controls were compared to the high iron group ( $p = 0.037$ ). These findings were supported by phantom experiments. These results suggest that there is a relationship between hepatic iron and increased L1 aBMD *Z-scores* in highly iron-loaded patients. Given patients with hemoglobinopathies are at increased risk for osteoporosis, clinicians should maintain a higher index of suspicion when diagnosing low bone mass.

### Keywords

Artifact; DXA; iron; vertebrae

### Introduction

In the last three decades, though scanner and detector technology have improved tremendously, the fundamental principle behind dual-energy X-ray absorptiometry (DXA)

\*Address correspondence to: Haven Allard, Children's Hospital Oakland Research Institute, 5700 Martin Luther King Jr. Way, Oakland, CA 94609, USA. hmallard@eckerd.edu.

has remained unchanged. DXA measures the attenuation of X-rays with high and low energy photons transmitted through the body from a single source (1). The X-rays that are used (typically 70 – 140 kV) in bone densitometers must be capable of passing through the body and table that the patient is lying on, and detected by sensors in the scanning arm, but also have a low enough X-ray emission dose to be used frequently for adult and pediatric diagnostic use. DXA employs the use of high and low energy X-rays, using the difference in the attenuation of tissue and bone, in order to calculate bone density. This is calculable because the composition of bone is made up primarily of calcium and phosphorus, elements that are much heavier with larger atomic numbers than those that constitute soft tissue: carbon, hydrogen, oxygen, and nitrogen. The two different levels of radiation emitted from DXA allow bone to be distinguished from soft tissue using a simplified two-compartment model, algebraic equations, and mass attenuation coefficients (see Fig. 1) for each X-ray energy (1).

One limitation of DXA technology is that images exist in 2-dimensional planes, therefore there is potential for artifacts to occur between the skeletal region of interest and the X-ray beam. Artifacts could affect the derived areal bone mineral density (aBMD) measurement. Under normal circumstances, soft tissue is subtracted from bone during the application of the manufacturer's analysis algorithm, allowing bone mineral density to successfully be measured without interference from surrounding tissue. Problems with successful DXA imaging and analysis arise when a patient's anatomy is altered, whether that be from external modifications (e.g. body piercings) or internal fixations (e.g. orthopedic hardware postsurgery), the anatomy no longer consists of simply soft tissue and bone. Studies have reported that artifacts such as dense clothing, surgical clips, jewelry, and other foreign objects can affect aBMD by DXA (2–5). External artifacts are typically removed prior to scanning, but internal artifacts are much more difficult to control and might interfere with DXA scans.

Ingested and/or absorbed metallic material has the potential to interfere with DXA imaging. One example of this artifact is barium sulfate, used for swallowing studies. Similar to barium, iron has a greater relative density and high atomic number, compared to elements in soft tissue. Therefore, in high concentrations iron deposited in soft tissue appears opaque, similar to bone. Iron can accumulate in extreme amounts in the liver,  $>9000 \mu\text{g Fe/g wet tissue}$  (54 mg/g dry weight) in patients receiving chronic red blood cell transfusions who are not adequately chelated. Patients with hemoglobinopathies are particularly at risk of developing systemic iron overload.

In our clinic we have been able to visualize the liver as an opaque region on whole body and vertebral scans by DXA in some of our severely iron-overloaded patients with sickle cell disease (SCD) and thalassemia (Thal) (see Fig. 2a and b). This opacity in the region of the liver led us to the observation that there is a strong, predictive relationship between high liver iron concentration and increased density by DXA (6). It is unknown if iron in hepatic tissue of highly iron-loaded patients could be mistakenly assessed by DXA as bone and if this would lead to increased aBMD lumbar spine Z-scores derived by DXA. The potential for this phenomenon to occur is a result of the DXA examination being performed on patients in the supine position. As a result, iron in the liver, which typically lies over lumbar

vertebrae L1/L2 could interfere with the scan of those overshadowed vertebrae. Our hypothesis is that iron in the liver of heavily iron-loaded patients will artificially raise aBMD in the spine, lead to an error and thereby to a potential misinterpretation of the DXA scan.

## Materials and Methods

A three-part study was designed to address our research question: (1) a retrospective chart review of patients with hemoglobinopathies (SCD and Thal) who visited UCSF Benioff Children's Hospital Oakland (BCHO) for both bone health (by DXA) and liver iron concentration (LIC) assessments on the same day; (2) assessment of healthy controls, with normal liver iron concentration, for comparison of DXA and LIC assessment; (3) and creation of artificially iron-loaded phantoms to further explore the relationship of known concentrations of iron with density assessments by DXA. The full study was approved by the Institutional Review Board at (BCHO).

## Inclusion/Exclusion Criteria

Data for the retrospective study were collected from patients with SCD and Thal who had a LIC measurement by biosusceptometry (SQUID: super-conducting quantum interference device) and bone health assessment by DXA completed on the same day at BCHO between January 2, 2002 and June 27, 2014. All lumbar scans were first reviewed by one investigator (HA), and excluded if artifacts were visible (e.g. surgical clips, gall stones, gastrostomy tubes, and navel piercings). Individual vertebral BMD *Z-scores* were derived, and comparisons made between the lumbar vertebrae that may be covered by the liver (L1 or L1/L2) with the *Z-scores* of L3 and L4 (see Fig. 3). Lumbar DXA scans were then matched with their corresponding LIC value measured on the same day. Data from all patient visits that qualified for inclusion were collected, even if a patient had two or more qualifying visits.

Healthy controls were assessed for eligibility to participate using a simple health questionnaire which asked questions related to their basic medical history and any potential nonremovable metal artifacts. To obtain a sample population representative of our heavily iron-loaded patient population, controls were selectively recruited by matching sex and age (within 5 yr) with our patients with an LIC >5000  $\mu\text{g Fe/g}$  wet weight. Although some of our patient population were younger than 18 yr old at the time of their bone and liver iron concentration assessments, only healthy controls >18 yr of age were allowed to participate in this study. Subjects were excluded if they were pregnant, had a body weight greater than 300 lbs (DXA table limit), wore nonremovable metal piercings, had metal dental work or metal implants, had a hemoglobinopathy or hemochromatosis, or a family history of hemoglobinopathy. Control subjects were also excluded if any lumbar vertebrae had an aBMD *Z-score* > +2.0 or < -2.0 or a LIC > 500  $\mu\text{g Fe/g}$  wet weight.

After subjects consented to the study, height and weight were measured and body mass index ( $\text{kg/m}^2$ ) calculated. All female subjects provided a urine sample before DXA scan for a pregnancy test. Subjects removed all clothing and jewelry that could potentially lead to an interference with the DXA or SQUID procedure. Subjects were then scanned according to

the same protocols followed when assessing bone density in our patient population, using the identical DXA imager and the same scan and analysis modes that were used on patients (Hologic Discovery A, Posterior-Anterior spine scan, Fast Array Mode). On the same day, controls were assessed for a liver iron concentration by SQUID.

There are no data available in the literature for which to base the sample size for the number of controls to recruit for this study. All retrospective data were gathered from patients which yielded ~60 patients visits with LIC >3000  $\mu\text{g Fe/g}$  wet tissue. A 2:1 ratio of cases to controls was calculated to be sufficient for 80% power at alpha 0.05 to observe a 1.0 standard deviation (SD) (1.0 Z-score) difference in L1 to L3/4 between cases and controls.

To address the third part of this study, a phantom was designed to mimic the geometry and iron concentration of a heavily iron-overloaded human liver. KNOX<sup>®</sup> brand gelatin and iron(II) sulfate heptahydrate were used in the preparation. The target phantom iron concentration was 9000  $\mu\text{g Fe/g}$  wet tissue. Vertebral DXA scans were taken with the phantom positioned obtusely overlying L1/L2 of the DXA daily quality control phantom to mimic the anatomical position of the liver in vivo (see Fig. 3a). The phantom was measured by DXA to model the potential effect of iron on measurement patterns by DXA. Twenty separate phantom scans were performed using the identical protocols followed when assessing both patients and control subjects (see Fig. 3b and c). Ten scans were taken without the liver to establish a control (see Fig. 3d) and ten scans were taken with the iron-loaded liver phantom (Fig. 3a). The phantom was also measured by SQUID to confirm iron concentration. All DXA scans were performed by one of two operators (EF or LC) using a Hologic Discovery A instrument (software version 12.6.1). The in vivo precision of DXA measurements was determined by duplicate measurements of 30 healthy subjects of similar age to those enrolled in this study. The root mean square error (RMSE) of spine aBMD performed at BCHO was 0.026  $\text{g/cm}^2$ ; spine BMC 1.86 g. The in vitro coefficient of variation (CV) of the DXA instrument was less than 1% for daily standard spine phantoms. LIC concentration was taken by SQUID biosusceptometer, model 5700 Ferritometer<sup>®</sup>(manufactured by Tristan Technologies, San Diego, CA).

## Data Analysis

Data were first plotted and tested for the assumption of normality while checking for outliers, ranges, and distribution assumptions. As expected, LIC was not normally distributed but was skewed to the right with a small percentage of patients with very high LIC and the majority of patients and controls with levels closer to the normal range. Summary statistics were then computed including means, SD and 95% confidence intervals (CI) for all variables within each group (Control, SCD, Thal, and LIC quartiles). Pearson's chi-square or Fisher's exact tests were used to assess differences in categorical variables (e.g. sex) between groups at the maximum time-point. For continuous variables, Student's *t*-tests were used for normally distributed data and Mann-Whitney tests for highly skewed data. For the purpose of this report, 'max time-point' was defined as the time-point during which the highest liver iron concentration was recorded for each case subject (SCD and Thal). LIC was analyzed in quartiles, with those with the lowest and highest quartiles compared first. The L1–L4 BMD *Z-scores* from the subjects in the highest and lowest LIC

groups were analyzed as unique data points using Student's *t*-tests. Paired *t*-tests were used to explore differences within subjects L1 vs L3/L4 BMD *Z*-scores for the controls and then the low-LIC and high-LIC groups separately. Similar paired *t*-test analyses were made for the iron phantom analyses. Following these cross-sectional analyses, the entire data set, which included multiple time-points per case subject, was analyzed. Multiple linear regression analyses were used to explore the relationship between the L1\_L3/4 *Z*-score difference observed at the spine and variables which may be associated with the variability in spine *Z*-score such as LIC, BMI, age, and PA (Posterior-Anterior) aBMD *Z*-score. Statistical analyses were conducted using Stata 9.2 (Stata, Inc., College Station, TX), and considered significant with a  $p < 0.05$ .

## Results

Thirty-one healthy controls were initially enrolled and assessed, 4 of which were excluded due to a LIC by SQUID  $>500 \mu\text{g Fe/g}$  wet tissue, and another 5 whom had lumbar vertebrae PA aBMD *Z*-scores  $> +2.0$  or  $< -2.0$ , therefore 22 controls are included in this analysis. Through retrospective chart review, patient data were extracted from a total of 114 patients. From these cases, a total of 248 scans were collected. This is after 24 scans and 6 patients were removed from the study due to artifacts obstructing scan analysis of one or more of the vertebrae of interest. Of the final patient population ( $n = 114$ ), 48 of them had more than one visit, and 24 patients had 4 or more combined visits (DXA + SQUID) during this 12yr study interval. All initial analyses were made using only the 'Maximum' LIC from each patient ( $n = 114$ ).

Patients in this study had a wide range of liver iron values, compared to healthy control subjects (Table 1), a difference which was significant ( $p < 0.001$ ). A normal range for LIC is typically considered to be between 80 and 370  $\mu\text{g Fe/g}$  wet tissue for children and 90 and 480  $\mu\text{g Fe/g}$  wet tissue for adults (7). There was no significant difference between LIC amongst the 2 patient groups given the large LIC range. As expected, spine PA aBMD *Z*-score was significantly lower in the patients with hemoglobinopathies vs the controls ( $p < 0.001$ ).

Initially, patient data was categorized into four groups (quartiles) based upon LIC (Group 4: LIC  $>4000 \mu\text{g Fe/g}$  wet tissue  $n = 30$ ; Group 3: LIC 4000-2201  $\mu\text{g Fe/g}$  wet tissue  $n = 29$ ; Group 2: LIC 2200-1001  $\mu\text{g Fe/g}$  wet tissue  $n = 38$ ; Group 1  $<1000 \mu\text{g Fe/g}$  wet tissue  $n = 17$ ). A *t*-test comparing Group 1 (lowest LIC) to Group 4 (highest LIC) revealed no significant difference in the difference of L1 and L3/4 ( $p = 0.287$ ). However, it became clear that the variability around the L1 to L3/4 difference for the highest group was too great (CV 118%), therefore we sought to find a more clinically significant LIC threshold to determine our cut-off groups. Therefore, groups were revised so that Group 4 was categorized as extreme iron-overload (see Table 2), and would consist of patients with a LIC  $>5000 \mu\text{g Fe/g}$  wet tissue (or 30 mg/g dry wt,  $n = 18$ ) and Group 3 (High LIC) consisted of patients with a LIC range of 2201-5000  $\mu\text{g Fe/g}$  wet tissue ( $n = 41$ ). With this adjustment, the CV around the L1 to L3/4 difference for the highest group was reduced to 78%, and a significant difference was observed between the patients with LIC  $<1000$  and those with extreme liver iron loading ( $>5000 \mu\text{g Fe/g}$  wet tissue, Fig. 4,  $p = 0.043$ ). The control group (LIC  $< 500$

mg Fe/g) was then compared to the extreme iron-overload Group 4 (LIC > 5000), again revealing a significant difference in the L1 to L3/4 Z-score ( $p = 0.037$ ).

The associations of sex, age, BMI and spine BMD Z-score were then explored in relation to the primary variable of interest L1\_L3/4 difference, using multiple linear regression models. Neither age nor sex were found to have a significant effect on the relationship between L1\_L3/4 difference and the grouping variable (LIC). However, both spine BMD Z-score and BMI had independent effects on the relationship between the dependent variable L1\_L3/4 difference, and our primary independent variable, LIC group (both  $p < 0.05$ ). Therefore, when LIC, spine BMD Z-score and BMI were all included in the same model, it became clear there was a significant interaction between the grouping variable (LIC), spine BMD Z-score, and BMI, so that LIC severity was no longer significant (Table 3). AP Spine BMD Z-score was negatively correlated with both LIC ( $p = 0.02$ ), as well as L1\_L3/4 difference ( $p = 0.03$ ).

These multivariate linear models were then expanded to include all the longitudinal data from each patient, including multiple visits and considering each visit as an independent data point ( $n = 274$  visits). Spine BMD Z-score and body mass index were strongly related to the L1\_L3/L4 difference, however, the severity of liver iron no longer contributed significantly to the model (Table 4).

## Phantom Data

Direct analysis of the heavily iron-loaded gelatin phantom by SQUID revealed that the phantom concentration was  $9553 \pm 428 \mu\text{g Fe/g}$  wet tissue. When this heavily iron-loaded phantom was scanned obtusely on the Hologic lumbar spine daily calibration phantom (Fig. 3a), there was a significant difference between the L1\_L3/4 Z-score difference of the Hologic Spine QC Phantom scanned without the iron-loaded gelatin phantom vs those scanned with the iron-loaded gelatin phantom (L1 to L3/4 SD Difference of: 1.0  $p < 0.001$ ).

## Discussion

This study revealed that there is a significant relationship between iron deposited in the liver and the accuracy of vertebral bone density measurements by DXA, particularly for patients with a liver iron concentration  $>5000 \mu\text{g Fe/g}$  wet tissue. The relationship is attenuated by a patient's aBMD Z-score, the lower the bone mass, the greater the potential 'masking' effect of hepatic iron. These findings are relevant as they illustrate the potential for an iron-overloaded liver to significantly alter the total spine aBMD Z-score by DXA, thereby leading to misinterpretation of bone densitometry scans in a patient receiving chronic red blood cell transfusions such as one with SCD or Thal. These results suggest that failing to correct for the iron contribution to the DXA spine scan in iron-overloaded patients could lead to under diagnosing low bone mass in at-risk patient populations.

In select patients, lumbar spine DXA scan is often the only available site for bone health assessment due to orthopedic interventions at the proximal hip (e.g. due to osteoarthritis or osteonecrosis particularly in SCD patients). Moreover, bone microarchitecture can now be derived from the vertebral site using the trabecular bone score software. This new tool has

been shown to add greater predictability to fracture risk than FRAX<sup>®</sup>, a commonly used diagnostic tool that assesses the 10 yr probability of fracture risk (8). For patients with Thalassemia, spine BMD is consistently lower than hip BMD Z-score and highly predictive of fracture (9,10). Thus, if vertebral scans are unreliable due to liver-iron interference, or affected and undetected by clinicians, then a diagnosis of low bone mass may be delayed or complicated. In this study, we estimated that 8% of patients with severe liver iron-overload would be misdiagnosed as having normal vertebral bone density, when in fact they have low bone mass ( $Z$ -score  $< 2.0$ ) that was masked by the liver iron artifact effect. Though this may seem clinically insignificant due to the small percentage of patients affected, if osteoporosis progresses undetected, within a few years fragility fracture may result, leaving patients debilitated with a significantly reduced quality of life.

Osteoporosis is a frequent complication in patients of hemoglobinopathies, such as SCD and Thal, because of the relationship between iron storage, bone development, and function. Specifically, patients can develop whole body iron overload due to chronic red blood cell transfusions or increased dietary iron absorption. The liver is the primary storage site for the body, but when iron within the liver becomes saturated, excess iron can “spillover” into the spleen, bone marrow, heart, and endocrine organs. Iron toxicity in the endocrine organs leads to diabetes, hypothyroidism, and hypogonadism, which places patients at increased risk for low bone mass and fracture (10,11). Some non-transfused thalassemia patients have ineffective erythropoiesis, resulting in hyperplasia (bone marrow expansion), which interferes with the structure of bone and causes a decrease in bone density to compensate for the larger area of bone marrow (12). This decrease in trabecular bone density is why it is important to monitor aBMD in patients, particularly those with Thal. Increased risk of fracture compounded with the potential for increased liver iron concentration to artificially increase aBMD  $Z$ -scores by DXA ultimately means that the clinician should maintain a higher index of suspicion when observing DXA scans and should understand this risk of interference in lumbar vertebral scans by DXA from this population of iron-overloaded SCD and Thal patients.

In the population that is represented in the study, low-trauma fracture risk becomes significant in survival and quality of life. While there is a paucity of literature on fracture risk in non-iron overloaded sickle cell disease patients, there are some data on the epidemiology of fracture in iron-overloaded patients with hemoglobinopathy and some data on non-loaded thalassemia patients. It is our hope that our results and discussion can be considered in conjunction with this data in order to increase clinician awareness of potential fracture risk in patients who may have inflated BMD  $Z$ -score values by DXA due to iron-overload. Figure 5.

It is possible that clinicians could simply remove L1 or L1 and L2 from analysis, as opposed to forgoing the lumbar scans, and combine these data with data taken from the hip. The possible exclusion of two or more vertebrae becomes problematic as studies have shown that the least dense vertebrae is not the best predictor for osteoporosis (13), meaning that it is important to use the entire distribution of L1 – L4 if possible. It would then be best if the clinician could understand a minimum LIC threshold at which there should be no interference in aBMD  $Z$ -score analysis by DXA so that L1 and L2 can be used as predictors

for osteoporosis in patients with hemoglobinopathies. As our results suggest, this threshold may be around 5000  $\mu\text{g Fe/g}$  wet tissue (30 mg Fe/g dry weight), as we observed no statistical effect of liver iron at that tissue iron concentration. Only on rare occasions (see Fig. 6a–c) were we able to visualize what is most likely high liver iron concentrations in the “undo view” function window of the DXA program. In these instances, soft tissue was deleted next to L1/2 during analysis. Because this happened with such a low frequency (<2% of scans) we were unable to determine an appropriate way to account for the additive or subtractive effects of iron loaded soft tissue mapped by DXA.

Studies, such as the one conducted by Morgan et al, have shown that artifacts lying next to a region of interest, such as the lumbar vertebra, may cause a subtractive phenomenon in DXA analysis leading to fictitiously decreased aBMD *Z-scores* being reported (3). This could mean that individuals with an iron loaded liver and anatomy where the liver is to the side of L1 and L2, as opposed to directly or obtusely over it, may actually demonstrate significantly lower L1 and L2 *Z-scores*. This could become a possibility in individuals with severe liver disease (cirrhosis) as the tissue could recede such that it no longer lies over L1 and L2 when the body is in the supine position. Alternatively, this could occur in patients with elevated iron tissue in the spleen.

In addition, Morgan’s study has shown that vertebra with lower aBMD have a greater potential to be affected by artifacts than do higher aBMD vertebra (3). With the significantly lower BMD found in our patient population than in our healthy controls, it is probable that there is an enhanced effect of high LIC on DXA scans in our patient population. More research should be done comparing the effects of liver iron concentration on patients with lower density vertebrae vs patients with higher density vertebrae.

## Limitations

Ultrasound is used to locate liver positioning when a patient is being assessed for liver iron concentration by SQUID, a similar technique could be useful to determine if the anatomy of each individual with iron overload is such that the liver is positioned over L1/L2 when laying supine prior to a DXA scan. Without this, we are limited to assumption and the rare ability to visualize the liver in our highly-loaded population. If the liver is heavily loaded but not positioned over L1/ L2 this may not cause an additive effect as we would assume, but instead may even lead to a reduced aBMD *Z-score*.

The possibility that iron has “spilled over” into reservoirs in the bone marrow suggests that our highly-loaded population may appear to have increased aBMD Lumbar *Z-scores* by DXA. This may negate any additive effect of an increased LIC and further manipulate the *Z-scores* such that the clinician may feel as though their patient’s bone density is higher than it is in reality. This may explain why in our multivariate model the relationship between LIC and L1\_L3/4 difference was no longer significant.

A final note, current Hologic software does not provide individual vertebrae BMD *Z-scores* for pediatric patients. Therefore, it would be difficult to develop exclusion criteria for iron-



overloaded patients less than 21 yr as individual vertebral contributions could not be compared. Only older software allows for this comparison (Hologic v. 12.6.1 and earlier).

## Conclusion

Hemoglobinopathy patients with liver iron-overload (>5000  $\mu\text{g}$  Fe/g wet weight) are at risk for under diagnosis of osteoporosis due to the potential for iron in the liver to spuriously elevate lumbar spine BMD *Z*-scores, L1 and L2. The combination of elevated risk of osteoporosis and fracture in patients with hemoglobiopathies and substantial liver iron argues that clinicians should maintain a higher index of suspicion when diagnosing low bone mass in these patients.

## Acknowledgment

The authors of this paper would like to recognize Dr. Frans Kuypers for providing assistance with this study as well as the NIH for providing support for this project (R25 HL125451).

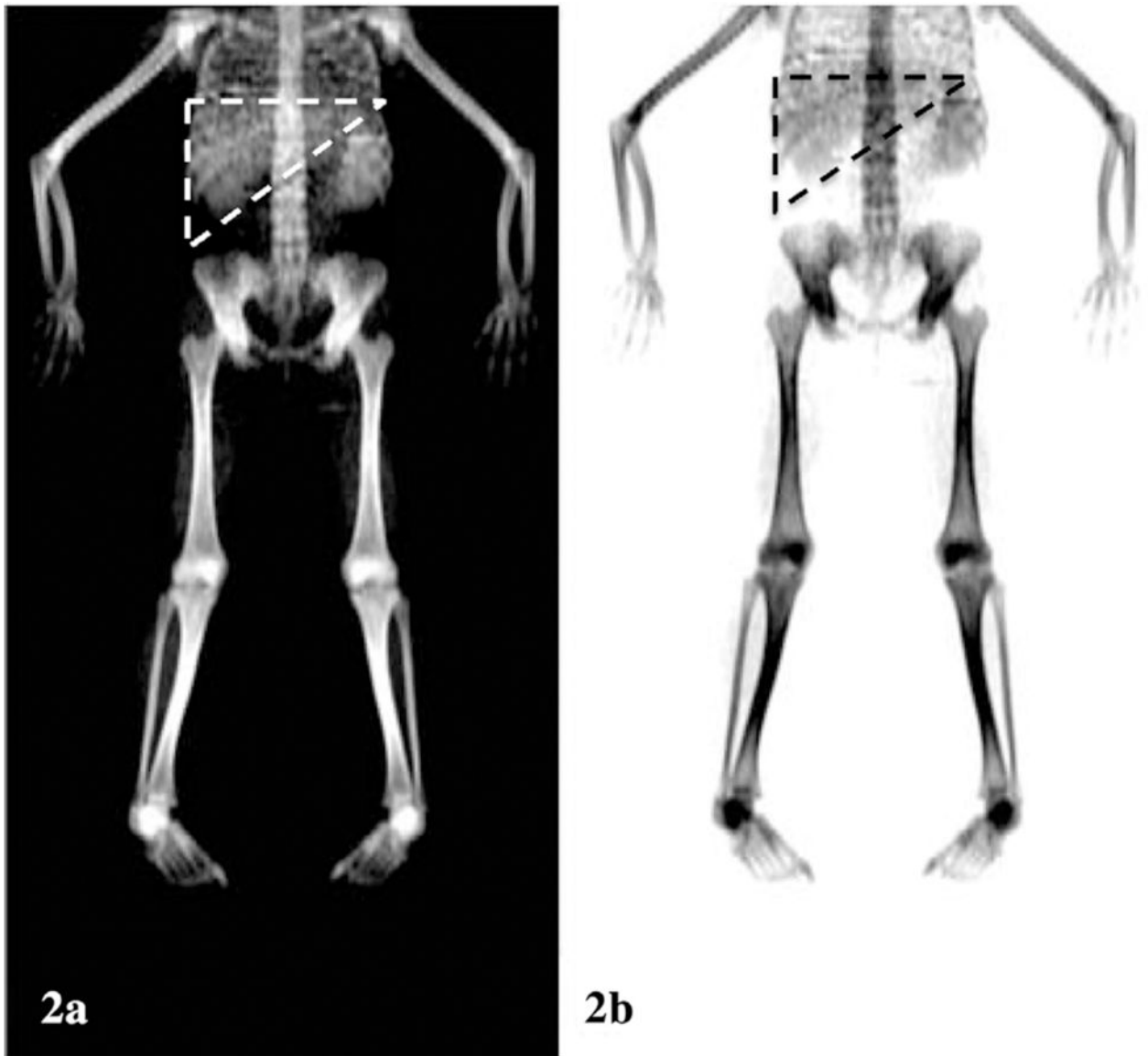
## References

1. Shepherd J, Crabtree NJ. 2016 Dual energy X-ray absorptiometry technology. In: Bone Health Assessment in Pediatrics, 2nd Eds. Fung EB, Bachrach LK, Sawyer AJ eds. Springer, Switzerland, 53-74.
2. Sherman ME, Shepard J, Frassetto L, Genant HK. 2002 Discrepancy in results between spine and hip scans of a woman with end stage renal disease. *J Clin Densitom* 5:9598.
3. Morgan SL, et al. 2008 "Black Hole Artifacts"—a new potential pitfall for DXA accuracy? *J Clin Densitom* 11:266275.
4. Lewiecki EM, et al. 2016 Best practices for dual-energy X-ray absorptiometry measurement and reporting: international society for clinical densitometry guidance. *J Clin Densitom* 19:127140.
5. Watts NB. 2004 Fundamentals and pitfalls of bone densitometry using dual-energy X-ray absorptiometry (DXA). *Osteoporosis Int* 15:847854.
6. Shepherd JA, et al. 2010 Dual-energy X-ray absorptiometry with serum ferritin predicts liver iron concentration and changes in concentration better than ferritin alone. *J Clin Densitom* 13(4):399–406. [PubMed: 20663700]
7. Nielsen P, Kordes U, Fischer R, Engelhardt R, Janka Ge. 2002 SQUID-biosusceptometry in iron overloaded patients with hematologic diseases. *Klin Padiatr* 214: 218–222. [PubMed: 12165905]
8. McCloskey EV, Oden A, Harvey NC, et al. 2016 A meta-analysis of trabecular bone score in fracture risk prediction and its relationships to FRAX. *J Bone Miner Res* 31 (5):940–948. [PubMed: 26498132]
9. Vogiatzi MG, et al. 2009 Bone disease in thalassemia: a frequent and still unresolved problem. *J Bone Miner Res* 24 (3):543–557. [PubMed: 18505376]
10. Fung EB, et al. 2008 Fracture prevalence and relationship to endocrinopathy in iron overloaded patients with sickle cell disease and thalassemia. *Bone* 43:162–168. [PubMed: 18430624]
11. Tubman VN, et al. 2015 Guidelines for the standard monitoring of patients with thalassemia: report of the thalassemia longitudinal cohort. *J Pediatr Hematol Oncol* 37 (3):162–169.
12. Brinker MR, et al. 1998 Bone mineral density of the lumbar spine and proximal femur is decreased in children with sickle cell anemia. *Am J Ortho* 27:4339.
13. Ryan PJ, Blake GM, Herd R, Parker J, Fogelman I. 1994 Distribution of bone mineral density in the lumbar spine in health and osteoporosis. *Osteoporosis Int* 4:6771.

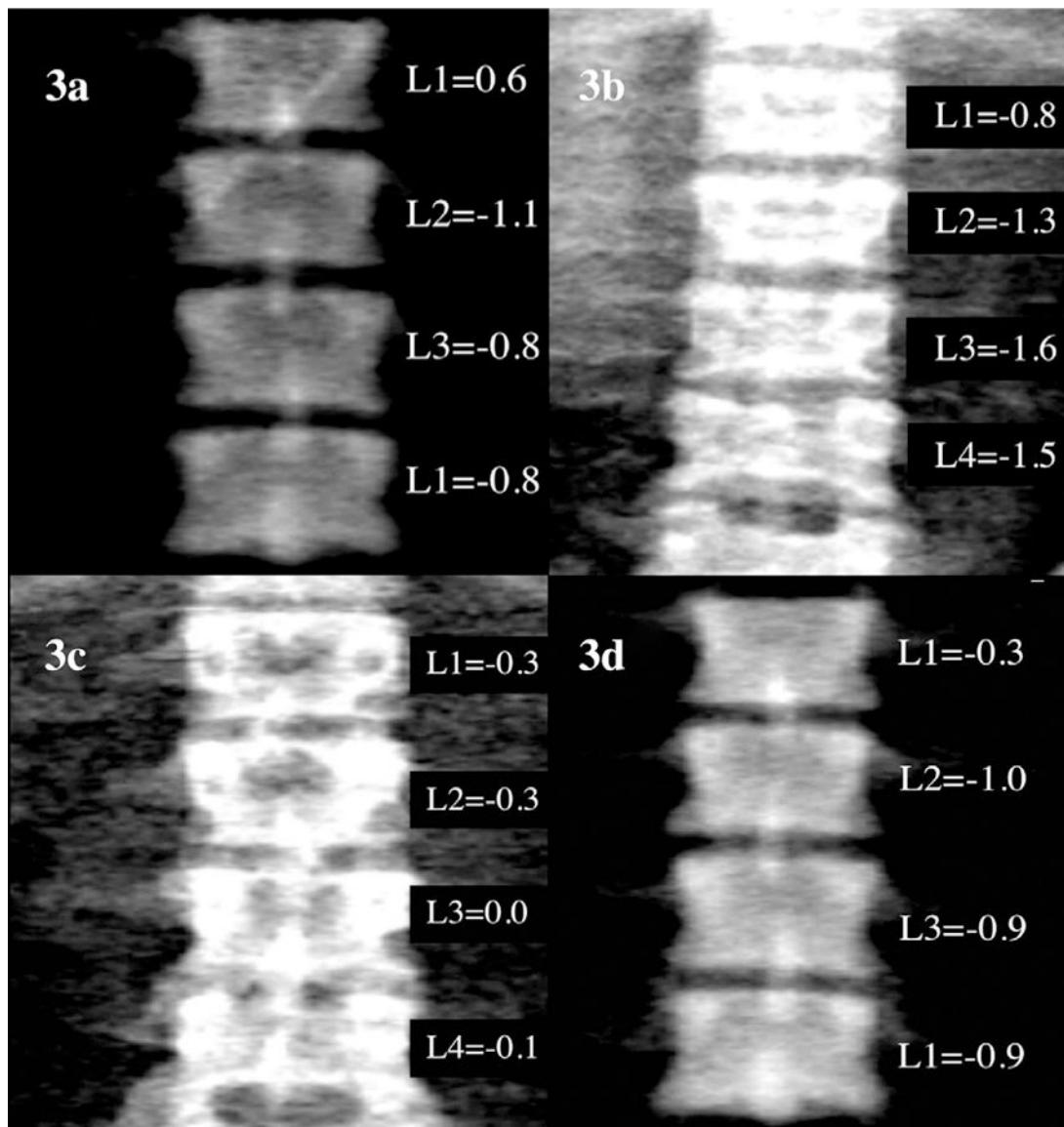
$$M_B = \frac{Ln(I_0^L / I^L) - kLn(I_0^H / I^H)}{\mu_B^L - k\mu_B^H}$$

**Fig. 1.**

The mass attenuation coefficient equation to extrapolate soft tissue for DXA.<sup>1</sup> Where B = bone, S = soft tissue, L = low energy photons, H = high energy photons,  $M_B$  = mass density of bone.

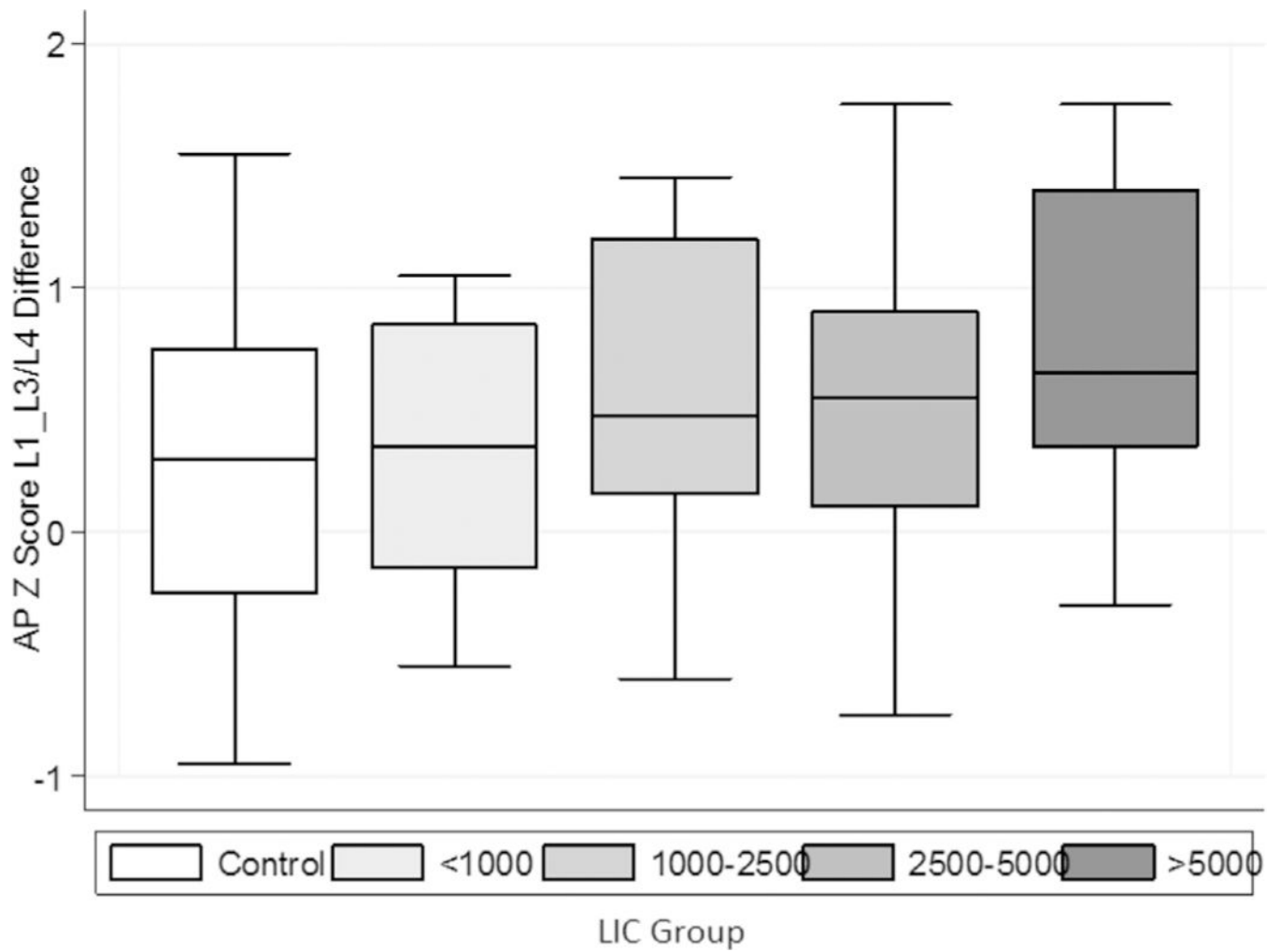


**Fig. 2.**  
(a/b) Whole body DXA inverse images displaying the severely dense region of both the liver (outlined in dashed lines) and spleen in a patient with thalassemia and LIC = 9898  $\mu\text{g Fe/g}$  wet tissue.

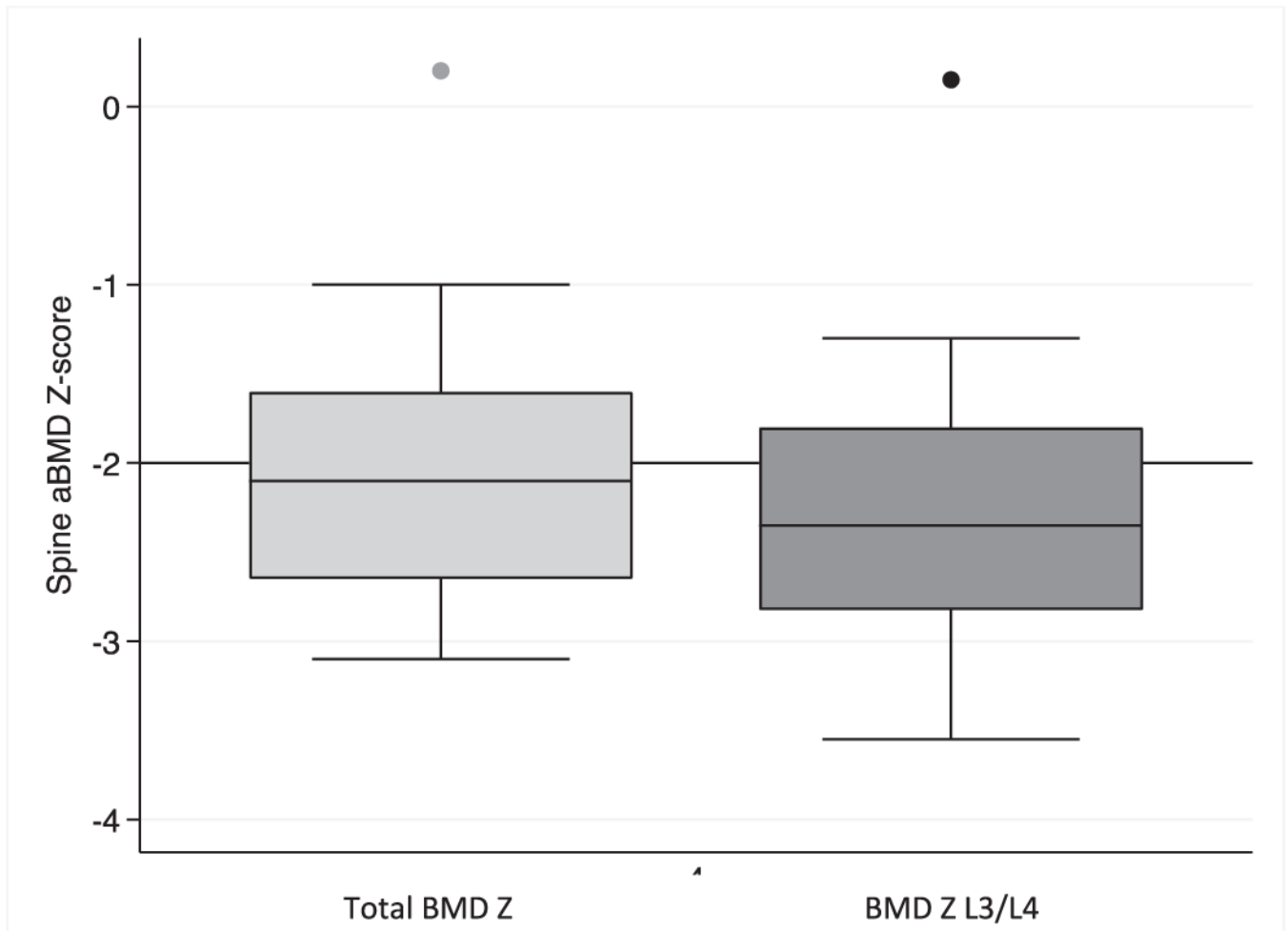


**Fig. 3.**

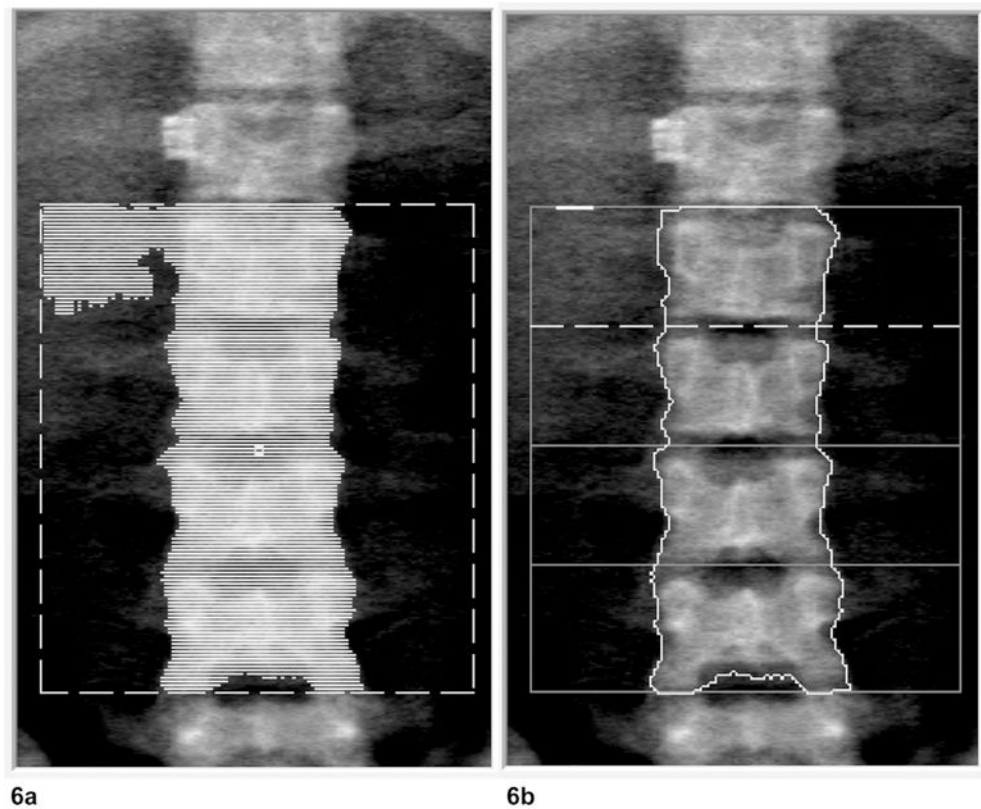
(a): Example of Hologic calibration phantom with artificially iron-loaded phantom ( $9000 \mu\text{g Fe/g wet wt}$ ) placed obtusely on top to simulate iron-loaded hepatic tissue. Scanned on Discovery A instrument as a patient (30 y.o., White, Male). Results for individual *Z-scores* provided. Difference between L1 and L3/L4 = 1.4 SD. (b): Example of a lumbar spine scan on a heavily iron-loaded patient with thalassemia (SQUID result:  $6877 \mu\text{g Fe/g wet wt}$ ). Difference between L1 and L3/L4 average = 0.7 SD. (c): Example of a lumbar spine scan from a healthy control subject (SQUID result:  $0 \mu\text{g Fe/g wet wt}$ ). Difference between L1 and L3/L4 average = 0.2 SD. (d): Example of Hologic calibration phantom as a control. Scanned on Discovery A instrument as a patient (30 y.o., White, Male). Results for individual *Z-scores* provided. Difference between L1 and L3/L4 = 0.6 SD.



**Fig. 4.** Vertebral aBMD L1 to L3/L4 Z-score difference by liver iron concentration group controls vs Group 4,  $p = 0.037$ ; Group 1 vs Group 4,  $p = 0.043$ . Group 1 <1000  $\mu\text{g Fe/g wt}$  ( $n = 17$ ); Group 2 1000-2500  $\mu\text{g Fe/g wt}$  ( $n = 38$ ); Group 3 2501-5000  $\mu\text{g Fe/g wt}$  ( $n = 41$ ); Group 4 > 5000  $\mu\text{g Fe/g wt}$  ( $n = 18$ ); LIC: liver iron concentration ( $\mu\text{g Fe/g wet tissue}$ ).



**Fig. 5.** Mean spine aBMD Z-score if all vertebrae are included (Light grey box plot, Mean: -1.9) compared to only BMD Z-scores for L3 and L4 (Dark grey box plot, Mean: -2.2): high iron-concentration sub-group 4 (LIC>5000  $\mu\text{g Fe/g}$  tissue). Overall: 8% of patients would be misdiagnosed as having normal bone mineral density if all vertebrae are used to classify vs only L3/L4.



Region	Area (cm <sup>2</sup> )	BMC (g)	BMD (g/cm <sup>2</sup> )	T- score	Z- score
L1	14.43	14.24	0.987		-0.2
L2	13.75	11.49	0.835		-2.4
L3	14.35	12.97	0.904		-1.8
L4	16.35	14.97	0.916		-2.1
TOTAL	58.88	53.67	0.911		-1.6

6c

**Fig. 6.**

(a): Example of lumbar scan with soft tissue mapped by DXA. 27 y.o. male, transfusion dependent thalassemia patient. LIC: 4891  $\mu\text{g/g}$  wet weight. Artifact: surgical clip at T-12.  
 (b): Example of lumbar scan with soft tissue mapped by DXA removed during analysis. (c): BMD Scores for L1 – 4 of example with mapped soft tissue after deleting mapped area next to L1.

Subject Demographics, Liver Iron Concentration and Lumbar Spine BMD by DXA in Subjects with Thalassemia, Sickle Cell Disease and Healthy Controls<sup>a</sup>

**Table 1**

	<b>Thal (n = 88)</b>	<b>SCD (n = 26)</b>	<b>Control (n = 22)</b>	<b>p value*</b>
Age (Mean ± SD)	24 ± 11.2	28.8 ± 15.1	29 ± 10.9	NS
Sex (Male/Female)	38 / 50	7 / 19	10 / 12	NS
BMI (Mean ± SD)	20.8 ± 3.7	21.4 ± 4.5	23.5 ± 3.6	NS
LIC (Mean ± SD)	2671 ± 1825	3585 ± 2357	147 ± 183	<0.001
LIC (Range)	524-9898	0-7829	0-485	<0.001
Lumbar spine aBMD Z-score (Mean ± SD)	-2.1 ± 1.25	-2.0 ± 1.32	-0.32 ± 0.65	<0.001

aBMD: posterior anterior areal bone mineral density; BMI: body mass index, kg/m<sup>2</sup>; LIC: liver iron concentration (µg Fe/g wet tissue); SCD: sickle cell disease, Age (years); Thal: thalassemia.

<sup>a</sup>Demographics for each subject group defined at the maximum time-point during which the 'highest' liver iron concentration was recorded for each case subject (sickle cell disease and thalassemia).

\* p value for differences between groups for categorical variables determined by Chi-square, and for continuous variables determined by ANOVA.



**Table 2**  
 Maximum Liver Iron Concentration Grouping Quartiles for Case Subjects with Hemoglobinopathies<sup>a</sup>

Iron group	n	Mean LIC	Std. Dev.	Min. LIC	Max. LIC
(1) Low LIC <sup>^</sup>	17	638	274	0	923
(2) Medium LIC	38	1627	332	1089	2198
(3) High LIC	41	3438	871	2214	4907
(4) Extreme LIC	18	6371	1388	5018	9898

<sup>^</sup> LIC, liver iron concentration (mg Fe/g wet tissue).

<sup>a</sup> Maximum for this report defined as the time-point during which the 'highest' liver iron concentration was recorded for each case subject (sickle cell disease and thalassemia).

**Table 3**

Regression Models on Maximum L1\_L3/4 Differences (n = 136).

	Coef.	S.E.	P >  t	[95% Conf. Interval]
<i>a. Univariate regression model:</i>				
Iron group	0.098	0.041	0.019	(0.016, 0.179)
<i>b. Multivariate regression model:</i>				
Iron group	0.078	0.042	0.068	(-0.005, 0.134)
aBMD Z - score	-0.089	0.040	0.030	(-0.170, -0.009)
BMI, kg/m <sup>2</sup>	0.029	0.013	0.026	(0.003, 0.055)

Dependent variable: L1 to L3/4 BMD Z-score difference.

Iron group, patient data categorized into four groups (quartiles) based upon LIC; aBMD, areal bone mineral density; BMI, body mass index, kg/m<sup>2</sup>; LIC, liver iron concentration ( $\mu\text{g Fe/g wet tissue}$ ); SCD, sickle cell disease. Age (years); Thal, thalassemia.

**Table 4**

Multivariate Regression Model including all Subject Time-points (n = 274) Dependent Variable: L1 to L3/4 BMD Z-score Difference.

	Coef.	S.E.	p-value	95% Conf. Interval]
LIC group	0.035	0.036	NS	(-0.035, 0.106)
aBMD Z - score	-0.141	0.028	0.000	(-0.195, -0.087)
BMI, kg/m <sup>2</sup>	0.031	0.009	0.001	(0.012, 0.049)

Iron group, patient data categorized into four groups (quartiles) based upon LIC; aBMD, areal bone mineral density; BMI, body mass index, kg/m<sup>2</sup>; LIC, liver iron concentration ( $\mu\text{g Fe/g wet tissue}$ ).