Intensity Standardization of Skeleton in Follow-Up Whole-Body MRI

Jakub Ceranka[1,2]([✉]), Sabrina Verga[1,3], Frédéric Lecouvet[4], Thierry Metens[5], Johan de Mey[6], and Jef Vandemeulebroucke[1,2]

[1] Department of Electronics and Informatics (ETRO),
Vrije Universiteit Brussel (VUB), Plainlaan 2, 1050 Brussels, Belgium
`jceranka@etrovub.be`
[2] imec, Kapeldreef 75, 3001 Leuven, Belgium
[3] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[4] Institut de Recherche Experimentale et Clinique (IREC),
Université catholique de Louvain, Louvain-la-Neuve, Belgium
[5] Department of Radiology, ULB-Hôpital Erasme,
Université Libre de Bruxelles (ULB), Brussels, Belgium
[6] Department of Radiology, Universitair Ziekenhuis Brussel, Brussels, Belgium

Abstract. The value of whole-body MRI is constantly growing and is currently employed in several bone pathologies including diagnosis and prognosis of multiple myeloma, musculoskeletal imaging and evaluation of treatment response assessment in bone metastases. Intra-patient follow-up MR images acquired over time do not only suffer from spatial misalignments caused by change in patient positioning and body composition, but also intensity inhomogeneities, making the absolute MR intensity values inherently non-comparable. The non-quantitative nature of whole-body MRI makes it difficult to derive reproducible measurement and limits the use of treatment response maps. In this work, we have investigated and compared the performance of several standardization algorithms for skeletal tissue in anatomical and diffusion-weighted whole-body MRI. The investigated method consists of two steps. First, the follow-up whole-body image is spatially registered to a baseline image using B-spline deformable registration. Secondly, an intensity standardization algorithm based on a histogram matching is applied to the follow-up image. Additionally, the use of a skeleton mask was introduced, in order to focus the accuracy of algorithms on a tissue of interest. A linear piecewise matching method using masked skeletal region showed a superior performance in comparison to the other evaluated intensity standardization methods. The proposed work helps to overcome the non-quantitative nature of whole-body MRI images, allowing for extraction of important image parameters, visualization of whole-body MR treatment response maps and assessment of severity of bone pathology based on MR intensity profile.

Keywords: Whole-body MRI · Intensity standardization · Skeletal imaging

1   Introduction

The value of whole-body magnetic resonance imaging (MRI) in skeletal imaging is constantly growing and is currently getting more interest in investigation of several bone pathologies, including diagnosis and prognosis of multiple myeloma [1], bone marrow in paediatric age [3], musculoskeletal imaging [6] and evaluation of treatment response assessment in bone metastases [2,8,11].

Due to its high resolution, whole-body coverage and high sensitivity MRI can provide excellent definition of anatomical structures and underlying skeletal pathologies. Additionally, in combination with a follow-up scan, it allows for monitoring of changes in patients body composition and disease involvement providing reliable treatment response assessment parameters (i.e. change in cancer volume, number of metastases) and image response maps.

Follow-up MR images acquired in the same scanner do not only suffer from spatial misalignments caused by different patient positioning and changes in patients' body composition over time, but also intensity inhomogeneities, making the absolute MR intensity values inherently non-comparable. Therefore, due to the non-quantitative nature of MRI, intensities cannot be compared from one acquisition to another making it impossible to derive reproducible intensity measurements containing interpretable information. Standardized images can not be displayed with fixed windows without the need of per-case adjustment. Additionally, they limit the use of treatment response maps only to quantitative MR modalities, such as MR apparent diffusion coefficient (ADC) calculated using diffusion-weighted images. In order to successfully compare a baseline and a follow-up whole-body scan, both limiting factors have to be overcome, usually via the means of image post-processing techniques.

Whereas, intra-patient whole-body spatial image misalignment can be compensated by image registration [6,17], inter-scan intensity inhomogeneities bring a challenging problem. In the literature, few authors have described different intensity standardization methods for MR images, however most of the work was done in the field of neuroimaging, limiting the application perspective to a very specific domain and much smaller field of view.

Nyúl et al. [10] proposed a linear piecewise method of matching image histograms of brain images. First, a number of intensity landmarks representing statistical points (percentiles, modes) are found in the reference and target image histogram. Secondly, both image landmarks are mapped on the common reference intensity space using a piecewise linear transform.

Robitaille et al. [14] proposed a method similar to Nuyl with a different landmark detection algorithm. The method incorporates tissue spatial intensity information derived from the segmentation image allowing for detection of more precise, tissue specific landmarks.

Jäger et al. [5] represented a group of multi-modal reference and target brain images as an n-D joint probability histogram. The next step involved deformable registration of obtained n-D histograms, which provided the deformation field matrix. The latter was used to standardize intensity inhomogeneities between the reference and target image stack. Additionally, a method was adapted for whole-body MRI images.

In this work, we propose an extension of existing intensity standardization methods maximizing the intensity similarity of skeletal structures in whole-body MRI together with an extensive quantitative evaluation. A strong validation criterion of mean absolute difference is introduced, allowing for direct quantification of intensity profile separation. The performance of the proposed algorithm was compared with the state-of-the-art methods.

2   Materials and Methods

The skeleton standardization methodology consist of two steps. First, the follow-up whole-body image is spatially registered to a baseline image. Accurate alignment of baseline and follow-up images improves the similarity of the intensity histograms limiting the influence of intra-scan anatomical differences. Additionally, it allows for the introduction of strong validation criteria based on voxelwise intensity comparison, such as the mean absolute intensity difference. Secondly, four different image intensity standardization methods were implemented and validated, aiming for equalization of skeleton intensity profiles.

2.1   Spatial Registration

In order to spatially align the baseline and the follow-up whole-body image and compensate for the aforementioned spatial misalignment, image registration was used.

Registration was performed in a pairwise manner, taking the baseline whole-body image as the reference image, f, and a follow-up image as a moving image, g. The aim was to solve an optimization problem finding a spatial transformation $\mathcal{T}$ over the parameters $\boldsymbol{\mu}$, according to the following equation:

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} \ \mathcal{C}_{\boldsymbol{x} \in \Omega}\Big(f(\boldsymbol{x}), g\big(\mathcal{T}_{\boldsymbol{\mu}}(\boldsymbol{x})\big) \Big). \tag{1}$$

In (1), the spatial coordinate $\boldsymbol{x}$ is taken from the overlapping region Ω, in which we assumed an intensity interpolation scheme for the discrete images f and g. The registration is guided by the minimization of the chosen cost function $\mathcal{C}$. Due to non-quantitative nature of the MRI before intensity standardization, a mutual information (MI) cost function [9] was used:

$$\mathcal{D}_{MI}(f, g(\mathcal{T}_{\boldsymbol{\mu}})) = -\sum_{a,b} p_{fg}(a,b) \log \frac{p_{fg}(a,b)}{p_f(a)p_g(b)}, \tag{2}$$

where p_{fg} is the joint probability density function (PDF) of the images f and g, and p_f and p_g are the marginalised PDFs for the respective images. a and b are the image intensity values.

Three stage multi-resolution image registration consisting of a rigid, affine and deformable B-Spline [15] deformation was implemented in the freeware software package `elastix` [7]. For a deformable step, a bending energy penalty (BEP) was used [15]. Detailed registration parameters are provided in Table 1.

Table 1. Parameters used in the spatial registration step.

Parameter	Spatial registration		
Transform	Rigid	Affine	B-Spline
Metric	MI	MI	MI, BEP
Number of resolutions	3	4	4
Image pyramid schedule	4 2 1	8 4 2 1	8 4 2 1
B-Spline grid spacing	-	-	4 2 1 1
Final B-Spline grid spacing (mm)	-	-	15 15 15
Number of histogram bins	32	32	32
Metric 1 weight	1	1	1
Metric 2 weight	0	0	10
Max iterations	2000	2000	2000
Sampler	Random	Random	Random
Number of samples	2048	2048	2048

The registration was driven by high resolution 3D T_1 whole-body image and the resulting transformation field was used to map other modalities of lower image quality, i.e. diffusion-weighted images.

2.2 Intensity Standardization

We compared 5 different intensity standardization algorithms with increasing complexity based on histogram matching principle.

Method 1. Linear Scaling: Target image is linearly scaled to match the intensity distribution in a reference image. Because of the signal intensity outliers, we use the intensity range up the the 99.9% intensity percentile, which according to the Eq. 3 gives:

$$I_{LS} = I_{Rmin} \frac{I_T - I_{Tmin}}{I_{ToutlierPerc} - I_{Tmin}} (I_{RoutlierPerc} - I_{Rmin}). \tag{3}$$

Here, we denote I_R, I_T and I_{LS} as the reference, target and linearly scaled output image. I_{Rmin}, I_{Tmin}, $I_{RoutlierPerc}$ and $I_{ToutlierPerc}$ are the minimum intensity values and 99.9% intensity outlier percentile values of the reference and target image, respectively. Additionally, all other compared methods were initialized from the linearly scaled result in order to roughly align image intensity profiles before allowing for standardization with more degrees of freedom. Experiment was performed, showing a benefit of initialization by linear scaling.

Method 2. Piecewise Linear Matching of Intensity Histograms: The method is implemented similar to [10], where the basic idea is to find a linear

piecewise mapping that deforms the follow-up image intensity histogram so that it matches a baseline image histogram using intensity landmarks. In the first step, five landmarks, L, representing intensity percentiles of the baseline and follow-up image are calculated. Here, a number of $n = 5$ evenly spaced percentile values was chosen, $L = [0, 20, 40, 60, 99.9]$. Second, a piecewise linear normalization is applied, mapping a follow-up image landmarks to corresponding baseline image landmarks, creating $n - 1$ linear and independent transformations, each between two landmarks (see Fig. 1).
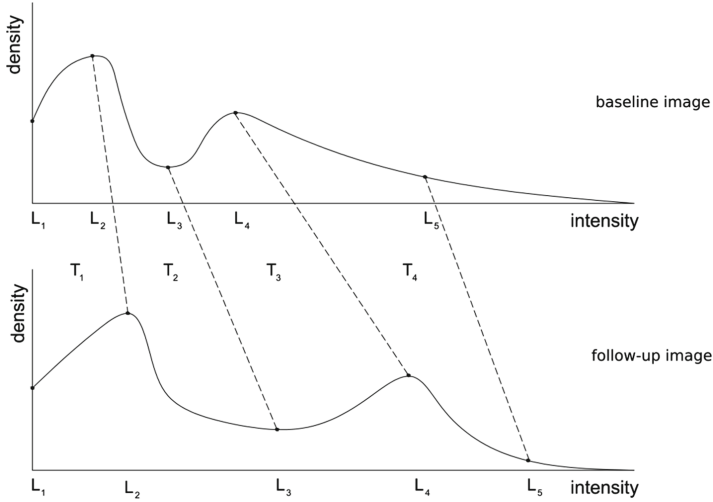

Fig. 1. Schematic representation of a linear piecewise transform. Two sets of landmarks L_{1-5} are detected in a reference and target image. Linear transformations T_{1-4} are used to standardize intensities between the images, mapping follow-up image intensities onto baseline image intensity profile.

Method 2.1. Piecewise Linear Matching of Masked Intensity Histograms: We propose a modification to the linear piecewise method by the introduction of the whole-body skeleton mask (see Sect. 2.3). Instead of taking all whole-body image voxels into account while calculating the intensity landmarks, only the masked tissues of interest will be used. Here, a 3D binary mask of the skeletal tissues is introduced, limiting excess of image information and focusing algorithm performance only on the chosen masked structure. Similar to method 2, five evenly spaced intensity percentiles were chosen as landmarks in the baseline and follow-up image, $L = [0, 20, 40, 60, 100]$, however, the range of intensities used was limited to the intensity range of the masked skeleton tissue. Later, as in method 2, piecewise normalization is performed taking into account updated landmark positions.

Method 3: Deformable Registration of Intensity Histograms: An image intensity histogram can be represented as a 1D image, where intensity values represent voxel count at each specific histogram bin. Therefore, the intensity standardization problem can be treated as a deformable image registration problem, aiming at finding a spatial transformation, $\mathcal{T}_{\mu}$, mapping a follow-up histogram image, $H(g)$, to a baseline intensity profile, $H(f)$, according to Eq. 1 (see Fig. 2). The resulting deformation field is used to correct intensities in the follow-up image [5]. Such method, gives more degrees of freedom compared to Method 2, allowing for smooth transformation and closer alignment of two intensity profiles. Here, a single-resolution deformable image registration with mean square difference cost function, bending energy penalty regularizer, histogram with 128 bins and a final B-Spline grid spacing of 30 pixels was used.
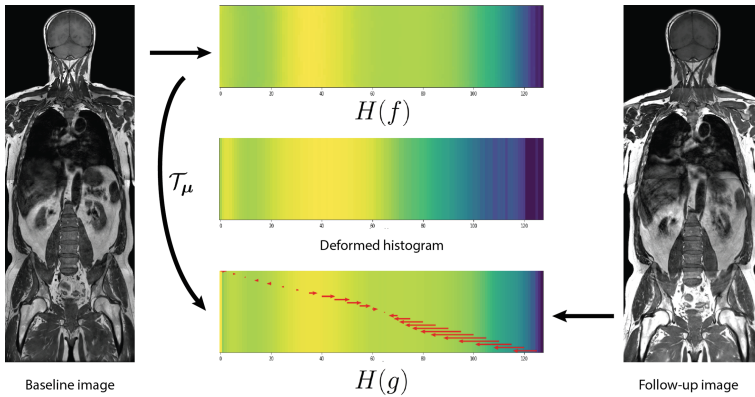

Fig. 2. Schematic representation of a deformable registration of two 1D histograms. The histogram of a follow-up image is deformably registered to the baseline image histogram. Obtained 1D deformation field (red arrows) is used to map intensities of the follow-up image onto baseline image intensity space. (Color figure online)

Method 3.1. Deformable Registration of Masked Intensity Histograms: The proposed method is a modification of method 3, where similarly to method 2.1, intensity histograms of the baseline and follow-up image are calculated only for the voxels included in the skeleton mask. Therefore, deformable registration is based only on intensities of interest, allowing for more precise intensity standardization transformation focusing on a chosen tissue of interest (i.e. bone).

2.3   Data Description

Experiments were performed on a 3D T_1 and diffusion-weighted whole-body images of prostate cancer patients with metastatic bone involvement and healthy

volunteers. Each patient had one follow-up examination, with an approximate 3–9 months between consecutive scans. The follow-up images of healthy volunteers were acquired during the same day, in a separate scanning session. 5 whole-body image pairs (baseline + follow-up) of the same subject consisting of 4 image station covering roughly head, torso, pelvis and legs were acquired. Images were obtained as a routine examination performed in the Cliniques Universitaires Saint-Luc, Brussels and Universitair Ziekenhuis Brussel. The study was approved by the Institutional Ethics Board of both institutions.

MRI: Whole-body stations were composed after independent image station preprocessing, which involved noise filtering using anisotropic diffusion followed by bias field correction [18] both implemented as a standard `Insight Segmentation and Registration Toolkit` (ITK) filters. Additionally, interstation intensity standardization was applied by scaling the intensity distribution of neighbouring stations to 99.9% intensity percentile based on the common station overlay region prior to the composition of the whole-body image from separate stations.

Anatomical whole-body image station were acquired as a T_1 weighted spin-echo sequence [12], with the following parameters: echo time (TE) = 8 ms, repetition time (TR) = 382 ms, matrix size of 480×480, pixel spacing 0.65 mm, slice thickness 1.19 mm. After the whole-body image reconstruction, spacing was equal to $1.2 \times 0.65 \times 0.65$ mm respectively in x, y and z direction with a matrix size of 210×1612–1705×768. Diffusion-weighted images were acquired with axial free breathing echo-planar DWI sequence (DWIBS) with a b-value equal to $1000 \, \text{s/mm}^2$. Following sequence parameters were used: TR = 8421 ms, TE = 66 ms, slice thickness 6.1 mm, matrix size 192×192, pixel spacing 2.3 mm, FOV = $440 \times 440 \, \text{mm}^2$.

Skeleton Segmentation Mask: For each whole-body image pair (baseline and follow-up image), the skeleton segmentation for a reference image was delineated using first, the *'GrowCutEffect'* application from *Slicer* [4] followed by manual refinement. Additional smoothing was applied using morphological operations. Aiming at the specific applications for bone pathologies (metastatic bone disease, multiple melanoma), only a selected number of bones with high probability of involvement were considered. This involved clavicle, spine from C2 vertebra to sacrum, pelvis and both femur bones. Tubercular bone as well as the cortical bone were included. Figure 3 illustrates the anatomical reference of the bones that are considered in this study together with corresponding manual segmentations.
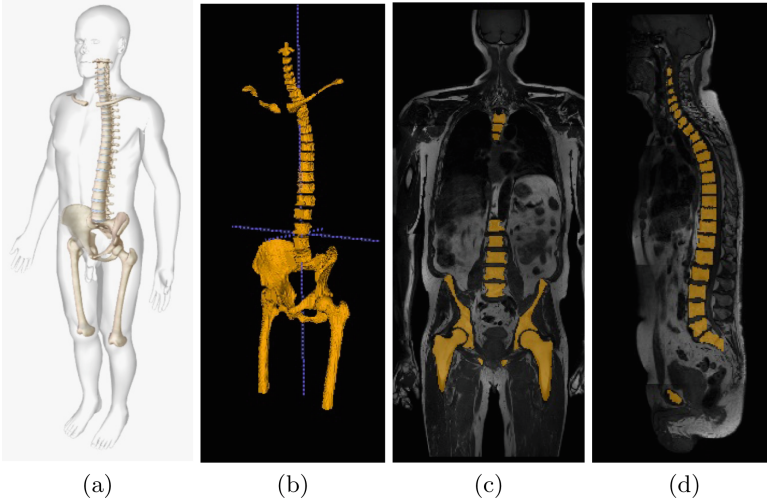
(a)   (b)   (c)   (d)

Fig. 3. (a) Anatomical reference from Bio Digital [13]. (b) Volume rendering from manual segmentation. (c) Coronal and (d) sagittal slice of a whole-body T_1 image in overlay with bone segmentation mask (yellow). (Color figure online)

2.4   Validation

Two validation criteria were used to asses similarity of skeletal intensity profiles between a reference baseline and target follow-up image.

Mean Absolute Difference: Corresponding voxel intensities were compared and summed into a mean absolute difference (MAD) value

$$MAD = \frac{\sum_x |f(x) - g'((\mathcal{T}_\mu(x))|}{N},$$
(4)

where, $f(x)$ and $g'(\mathcal{T}_\mu(x))$ are image intensities of the reference and spatially registered - intensity standardized target image in the corresponding voxel location x and N is a number of image voxels.

Kullback-Leibler Divergence: We have implemented the Kullback-Leibler divergence (KL) representing a distance measure between two discrete probability distributions (histograms)

$$KL_D = \sum_i P(i) log \frac{Q(i)}{P(i)},$$
(5)

where, $P(i)$ and $Q(i)$ are discrete probability distributions of a reference and standardized image at histogram bin i.

We can assume that if different tissue classes cover the same intensity range in both volumes, the histograms of a reference and target whole-body image will be as similar as possible, representing KL value close to zero.

Since not all of the data proved to be normally distributed ($p > 0.05$, Shapiro-Wilk Normality Test [16]), the Wilcoxon two-tailed, signed-rank test was used to investigate statistical significance of differences in validation criteria values between the non-standardized image and each of the registration strategies separately. The p-value used for the statistical significance test was equal to 0.05.

3   Results

All proposed intensity standardization methods were quantitatively validated and compared to a spatially registered and non-standardized whole-body image pair, representing a baseline value. Results of the validation criteria representing intensity standardization performance between baseline whole-body image and follow-up whole-body image, averaged over all subjects used, are presented in Table 2. Figure 4 shows the influence of the spatial registration and intensity standardization on baseline and follow-up image skeleton similarity. Figure 5 shows whole-body T_1 and DWI baseline and follow-up images before and after intensity standardization displayed with the same window and level setting. A sample T_1 functional response map indicating metastatic bone disease progression is shown in Fig. 6.

Table 2. Evaluation metrics averaged over 10 whole-body image pairs of T_1 and DWI modalities for the proposed methods ($\pm$ standard deviation). The best performing strategy in terms of average for each criteria is highlighted in bold. Statistical significance for each registration strategy and evaluation criterion, when compared to unregistered raw images is marked with an asterix (*).

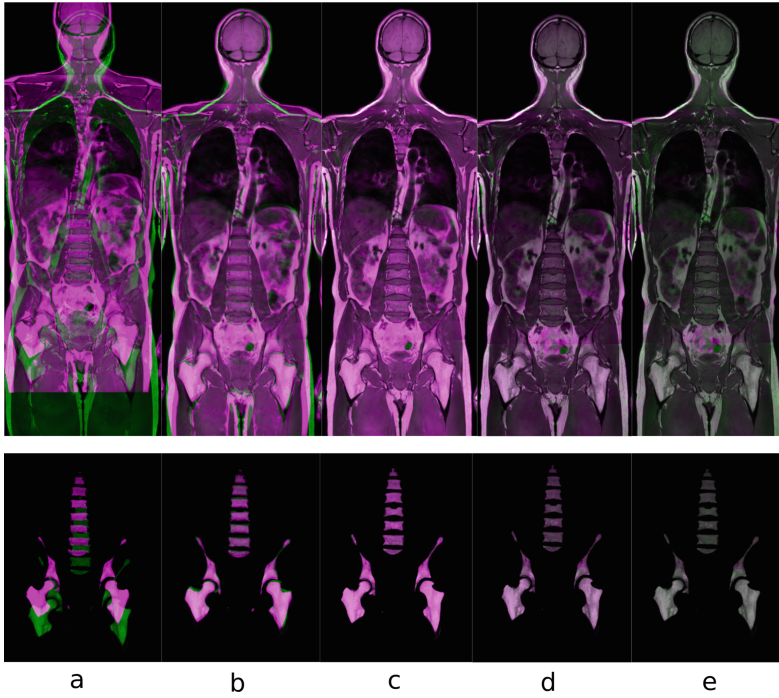|	MAD	KL divergence
No standardization	143.20 $\pm$ 126.67	0.540 $\pm$ 0.520
Method 1	76.15 $\pm$ 91.49	0.540 $\pm$ 0.520
Method 2	69.27 $\pm$ 81.77	0.468 $\pm$ 0.678
Method 2.1	**49.29 $\pm$ 50.68***	**0.095 $\pm$ 0.122***
Method 3	64.86 $\pm$ 60.22*	0.165 $\pm$ 0.169*
Method 3.1	64.32 $\pm$ 74.37*	0.484 $\pm$ 0.399

Fig. 4. Coronal overlay view of whole-body T_1 image (top) with extracted skeleton (bottom). Pink and green colours indicate intensity difference. **(a)** Raw images, **(b)** result after rigid registration, **(c)** result after deformable registration, **(d)** result after deformable registration with linear scaling of intensities (Method 1), **(e)** result after deformable registration with intensity standardization (Method 2.1). (Color figure online)

3.1   Computation Times

Processing was performed using a 2.5 GHz Intel® Core® i7-4870HQ processor and 16 GB RAM. Spatial registration inducing preprocessing steps and image re-sampling took around 30 min for an image pair. The entire standardization procedure (single threaded execution) for method 1 and all variations of method 2, took around 1 min. Method 3 with an execution time equal to 30 min, is considerably more expensive due to the deformable histogram registration and a higher number of intensity transformations equal to the size of the 1D deformation field.
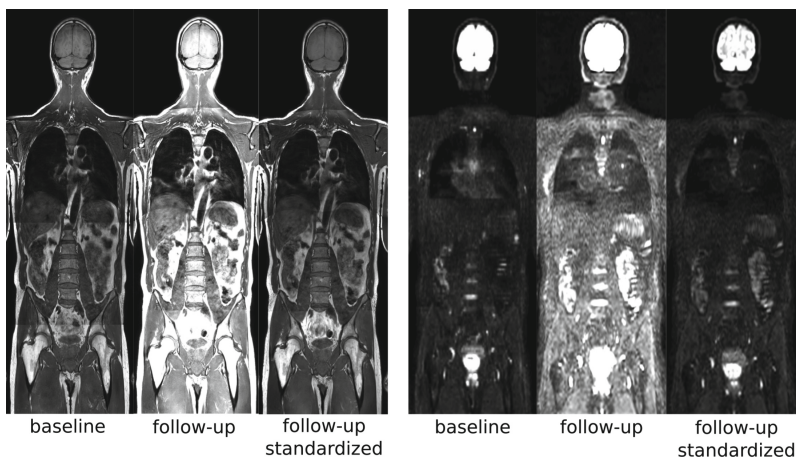
Fig. 5. Whole-body T_1 (left) and DWI (right) baseline and follow-up images before and after intensity standardization (Method 2.1). Images have been spatially registered. All images are displayed with the same window and level setting.
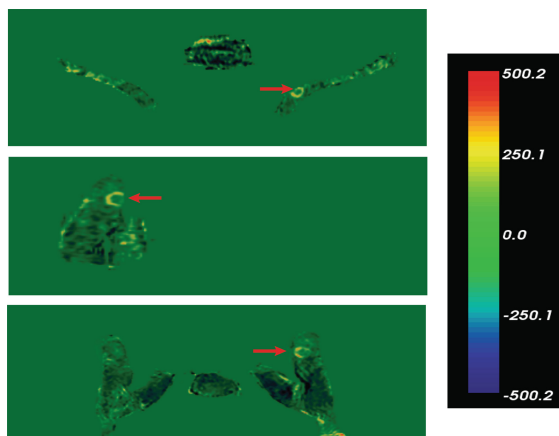

Fig. 6. From top to bottom: axial, sagittal and coronal view of functional response map calculated on T_1 intensity standardized image showing left upper pelvis with a visible progression of focal bone metastasis (red arrow). (Color figure online)

4   Discussion and Conclusion

In this work we investigated several strategies for intra-patient whole-body intensity standardization of skeleton profiles. Five different intensity standardization methods were compared and their performance was validated. Additionally, the use of spatial registration between the baseline and follow-up volumes, allowed for the introduction of strong validation criterion based on direct intensity difference - mean absolute difference of skeleton intensity profiles. The piecewise

linear method using the masked tissue of interest (Method 2.1) performed better than other evaluated methods, showing high stability and robustness of performance. Slightly worse performance of masked 1D deformable method (Method 3.1) might be caused by the limited amount of image information which corrupts the performance of deformable registration algorithms and the over-fitting of the match of the intensity profiles. Intensity standardization algorithms can be applied to any other tissue of interest if a specific mask representing a tissue type is provided.

Accurate intensity standardization of intra-patient MRI whole-body skeleton profiles, opens opportunities for whole-body quantitative follow-up, cohort comparison studies and functional response maps for non-quantitative modalities, considerably simplifying extraction of relevant quantitative information for healthy and disease.

References

1. Baur-Melnyk, A., Buhmann, S., Dürr, H., Reiser, M.: Role of MRI for the diagnosis and prognosis of multiple myeloma. Eur. J. Radiol. **55**(1), 56–63 (2005)
2. Blackledge, M.D., et al.: Assessment of treatment response by total tumor volume and global apparent diffusion coefficient using diffusion-weighted MRI in patients with metastatic bone disease: a feasibility study. PLoS ONE **9**(4), e91779 (2014)
3. Darge, K., Jaramillo, D., Siegel, M.J.: Whole-body MRI in children: current status and future applications. Eur. J. Radiol. **68**(2), 289–298 (2008)
4. Fedorov, A., et al.: 3D slicer as an image computing platform for the quantitative imaging network. Magn. Reson. Imaging **30**(9), 1323–1341 (2012)
5. Jäger, F., Hornegger, J.: Nonrigid registration of joint histograms for intensity standardization in magnetic resonance imaging. IEEE Trans. Med. Imaging **28**(1), 137–150 (2009)
6. Karlsson, A., et al.: Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. J. Magn. Reson. Imaging **41**(6), 1558–1569 (2015)
7. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging **29**(1), 196–205 (2010)
8. Lecouvet, F.E., et al.: Can whole-body magnetic resonance imaging with diffusion-weighted imaging replace Tc 99m bone scanning and computed tomography for single-step detection of metastases in patients with high-risk prostate cancer? Eur. Urol. **62**(1), 68–75 (2012)
9. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imaging **16**(2), 187–198 (1997)
10. Nyúl, L.G., Udupa, J.K., Zhang, X.: New variants of a method of MRI scale standardization. IEEE Trans. Med. Imaging **19**(2), 143–150 (2000)
11. Padhani, A.R., Makris, A., Gall, P., Collins, D.J., Tunariu, N., de Bono, J.S.: Therapy monitoring of skeletal metastases with whole-body diffusion MRI. J. Magn. Reson. Imaging **39**(5), 1049–1078 (2014)
12. Pasoglou, V., et al.: Whole-body 3D T1-weighted MR imaging in patients with prostate cancer: feasibility and evaluation in screening for metastatic disease. Radiology **275**(1), 155–166 (2014)

13. Qualter, J., et al.: The biodigital human: a web-based 3D platform for medical visualization and education. Stud. Health Technol. Inform. **173**, 359–361 (2012)
14. Robitaille, N., Mouiha, A., Crépeault, B., Valdivia, F., Duchesne, S.: Tissue-based MRI intensity standardization: application to multicentric datasets. J. Biomed. Imaging **2012**, 4 (2012)
15. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging **18**(8), 712–721 (1999)
16. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika **52**(3/4), 591–611 (1965)
17. Strand, R., et al.: A concept for holistic whole body MRI data analysis Imiomics. PloS one **12**(2), e0169966 (2017)
18. Tustison, N.J., et al.: N4ITK: improved N3 bias correction. IEEE Trans. Med. Imaging **29**(6), 1310–1320 (2010)