

Predicting Future Bone Infiltration Patterns in Multiple Myeloma

Roxane Licandro^{1,2}, Johannes Hofmanninger², Marc-André Weber³, Bjoern Menze⁴, Georg Langs²
`roxane.licandro@meduniwien.ac.at`

¹Visual Computing and Human-centered Technologies - Computer Vision, TU Wien

²Department of Biomedical Imaging and Image-guided Therapy - Computational Imaging Research Lab, Medical University of Vienna

³Institute of Diagnostic and Interventional Radiology, University Medical Center Rostock

⁴Institute of Biomedical Engineering - Image-based biomedical modelling, Technische Universität München

Abstract. Multiple Myeloma (MM) is a bone marrow malignancy affecting the generation pathway of plasma cells and B-lymphocytes. It results in their uncontrolled proliferation and malignant transformation and ultimately can lead to osteolytic lesions first visible in MRI. The earliest possible reliable detection of these lesions is critical, since they are a prime marker of disease advance and a trigger for treatment. However, their detection is difficult. Here, we present and evaluate a methodology to predict future lesion emergence based on T1 weighted Magnetic Resonance Imaging (MRI) patch data. We train a predictor to identify early signatures of emerging lesions before they reach thresholds for reporting. The algorithm proposed uses longitudinal training data, and visualises high-risk locations in the bone structure.

1 Introduction

Multiple Myeloma (MM) accounts for 10% of all bone marrow malignancies with an incidence rate of 6/100000 per year in western countries [5]. It is the second most common blood affecting malignancy, which disturbs the generation pathway of plasma cells and B-lymphocytes. Consequently, these cells proliferate uncontrolled and are transformed in a malignant way [10]. In addition, the production of large amounts of non-functional monoclonal antibodies is enforced, which affects the function of kidneys, increases the deficiency in immune response and in an advanced stage, influences the generation of bone forming and resorbing cells. MM starts at a precursor state of *Monoclonal Gammopathy of Undertmined Significance* (MGUS) and further evolves to an asymptomatic form of the disease *smoldering Multiple Myeloma* (sMM) with a predictable progression to the symptomatic form of MM [4].

Longitudinal bone infiltration patterns of MM progression The increased amount of plasma cells in MM leads to the alteration of bone remodelling mechanisms, by promoting bone resorption and inhibiting bone formation [10]. This results

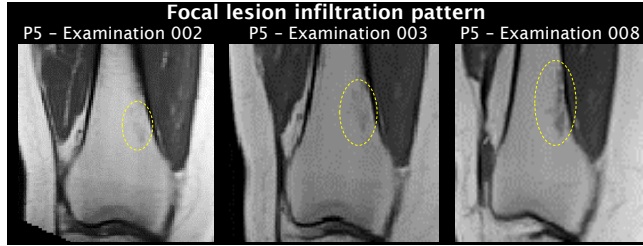


Fig. 1. Visualisation of an infiltration pattern of a focal lesion (yellow) using T1 weighted MRI scans over multiple examination time points of one patient.

first in the formation of focal or diffuse bone marrow infiltration. The gold standard for observing these initial infiltration patterns is MRI (T1, T2) [1][8][4]. Subsequently, the progression of the disease leads to the building of osseous destructions, which are observable using low-dose Computer Tomography (CT) [6]. Figure 1 illustrates the infiltration pattern of a focal lesion evolving at the distal part of the femur over three examination time points of a single patient.

Challenges Challenges of tracking lesions over time are, identifying early signatures of their emergence, accurate alignment of subject whole body images, imaging artefacts, and subtle non rigid deformations, as well as capturing the heterogeneity of diffuse infiltration patterns and their imaging signatures. Different treatment strategies and patient specific treatment response, and progression speed cause further variability. According to the results reported in the recent study of Mateos et al. [7], it is particularly important to assess high-risk sMM patients for developing MM and corresponding infiltration patterns, since a benefit for the patient from early therapy is observable.

Contribution Here, we propose and evaluate a predictor for future bone infiltration patterns. The algorithm uses longitudinal data to learn a local predictor of lesion emergence and change. We assess longitudinal relationships between subsequent stages of bone lesions and corresponding infiltration patterns of MM patients to provide a predictive signature for bone lesions. The contribution of the paper is three fold: (1) the longitudinal alignment of multiple bodyparts in whole body MRIs, (2) a classifier incorporating data from different disease stages in MM and (3) a local lesion risk score (LRS) to identify bone regions with a higher probability to evolve to diffuse or osteolytic lesions. We first give an overview of methodology and the data in Section 2. The evaluation results are presented in Section 3 and the conclusion of this work and possibilities for future work are summarized in Section 4.

2 Methodology

We first describe longitudinal alignment, and then introduce the methodology to estimate a local Lesion Risk Score (LRS) for future lesion emergence. Figure

2 illustrates the computation pipeline of the lesion risk score. It consists of four

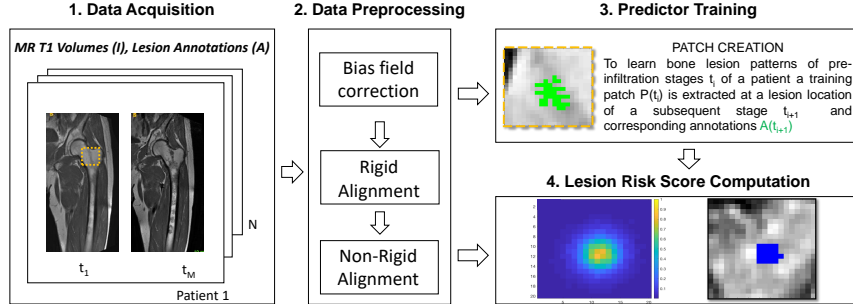


Fig. 2. Lesion Risk Score computation pipeline

components: data acquisition, data preprocessing, train and test data design and LRS computation. Details regarding the train and test data design are summarized in Section 3.

Longitudinal Alignment To perform subject specific longitudinal analysis of subsequent lesion states, we first register a patient’s data $I_{t_i} = \{I_{t_1}, \dots, I_{t_M}\}$. A patient’s image at a timepoint t_i is aligned to all subsequent timepoints $x = t_{i+1}, \dots, t_M$, depending on the number of available data. Bias field correction is used to process imaging data before alignment using FAST¹ [3] integrated in the FMRIB Software Library (FSL)². The aligned image $\overline{I_{t_i}(x)}$ is obtained following a two step registration procedure (cf. Equation 1).

$$\overline{I_{t_i}(x)} = I_{t_i} \circ \phi_{NR}((A * I_{t_i}), I_x) \quad (1)$$

Although longitudinal data is registered, patients’ shape vary over time, since the median interval between MRIs is 13 months. Also the acquired images do not visualise exactly the same body part snippets. Thus, we decided (based on experimental results), to use an affine approach first. This affine alignment A is performed using a block matching method for global registration. The resulting transformed image $(A * I_{t_1})$ is used as moving image in the second alignment step. Subsequently, a non-rigid deformation ϕ_{NR} to the target at time point x is estimated. For affine registration the function *reg.aladin* and for non rigid alignment the function *reg.f3d* are used, which are integrated in the NiftyReg toolbox³ [9]. For assessing the quality of registration, moving and target image were inspected manually using overlay visualisations and evaluated regarding correspondence of lesions’ position between the different time points.

¹ <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FAST> [accessed 11th of June 2018]

² <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL> [accessed 11th of June 2018]

³ <http://cmictig.cs.ucl.ac.uk/research/software/software-nifty/niftyreg> [accessed 11th of June 2018]

Patch Creation and Data Augmentation After alignment, MR images $\overline{I_{t_1}(x)}$ of a patient correspond to timepoints x and corresponding annotations S_x of the lesion in a future image of the patient. After the longitudinal alignment of acquisitions of one patient, datasets are created by using the intensity image of a subject at time point t and to this time point aligned annotations of lesions of subsequent time points $t + i$, where i is $1, \dots, n$ (n... number of available subsequent time points). Given those data pairs, patches are created in the lesion regions, by first computing the barycentre of the lesion’s annotation in the subsequent state and randomly moving a clipping window around it to avoid the predictor to learn uniform lesion positions. Additionally, we repeated this procedure for rotated intensity images and corresponding annotations. The rotations were performed in 20 degree steps.

Local Lesion Risk Score The proposed local Lesion Risk Score (LRS) uses early signatures of emerging bone lesions to predict future lesions and mark corresponding high risk locations. We use the computed pairs of image patches and corresponding lesion annotations of a subsequent state to train a random forest classifier that predicts future lesion labels from the present image patch data. During application, a score is obtained for each voxel position V by the probability predicted by the trained random forest for a new input patch.

$$LRS_V = P_{RF}(\overline{I_{t_i}(x)}) \quad (2)$$

We used the Python framework *sklearn* with an integrated Random Forest predictor⁴ with 10 decision trees and the following parametrisation: `n_estimators=10`, `criterion=gini`, `max_depth=2`, `min_samples_split=2`, `min_samples_leaf=1`, `min_weight_fraction_leaf=0.0`, `max_features=auto`, `max_leaf_nodes=None`, `min_impurity_decrease=0.0`, `min_impurity_split=None`, `bootstrap=True`, `oob_score=False`, `n_jobs=1`, `random_state=None`, `verbose=0`, `warm_start=False`, `class_weight=None` .

3 Results

In this section an overview of the evaluation dataset is given and the quantitative as well qualitative LRS evaluation results are presented and discussed.

Dataset In this study 220 longitudinal whole body (wb) MRIs from 63 patients with smoldering multiple myeloma (following the 2003 guidelines [2]) were acquired between 2004 and 2011. At least one wbMRI was performed per patient. The annotation of focal lesions is performed manually by medical experts starting at a lesion size bigger than 5 mm [1], since according to the IMWG consensus statement, from this size on, patients are considered to have symptomatic myeloma with therapy requirement. Table 1 summarizes the study participant’s

Table 1. Participants’ demographics

Patients	63 (24 female)
Therapy	Radiation or resection
Median age at initial MRI (yrs)	55
Age range (yrs)	29 -76
Median interval between MRIs	13 months
Median observation time	46 months

demographics. The protocol of this study was approved by the institutional ethics committee and all subjects gave their informed consent prior to inclusion. The scanning was performed on a 1.5 Tesla Magnetom Avanto (Siemens Healthineers, Erlangen, Germany) scanner. For the T1 weighted acquisition a turbo spine echo sequence (repetition time (TR): 627 milliseconds (ms), echo time (TE): 11 ms, section thickness (ST): 5mm, acquisition time (TA): 2:45 min) was performed of the head, thorax, abdomen, pelvis and legs using a coronal orientation and for the spine in sagittal orientation. No contrast medium was given. The duration of a scan was approximately 40 minutes long.

Train and Test Data Design In this study we used acquisitions from two body regions: In *region 1* thorax, abdomen and pelvis are visualised and in *region 2* the lower part of the pelvis, femurs and knees. These areas are considered since most lesions occur in those. In this work we observe two types of lesions and evaluate the performance of the methodology proposed separately for every type, with corresponding train and test sets: lesions which are *emerging*, i.e. which are not reported in the first scan, but in the subsequent scan, and *changing* lesions, which are annotated at both observed examination time points. For every patient we extracted image patches at lesion regions longitudinally over subsequent states of three different sizes ($10 \times 4 \times 10$, $20 \times 4 \times 20$ and $30 \times 4 \times 30$ voxels with a voxelspacing of $1.302 \text{ mm} \times 6 \text{ mm} \times 1.302 \text{ mm}$). To obtain a higher number of patches for the predictor training, data augmentation is performed resulting in *18 different patches per lesion*. To summarize, for emerging lesions we obtain 720 patches for region 1 and 504 patches for region 2 and for growing lesions we created 1026 patches for region 1 and 810 patches for region 2. Crossvalidation is used to generate test and training datasets, where a testset consists of 18 patches of a single lesion including the volumes of different orientation, which results in 40 folds for emerging lesions in region 1 and 28 in region 2 and 57 folds for growing lesion in region 1 and 45 in region 2.

Evaluation Setup For the quantitative evaluation and for obtaining comparability between the different tested setups, the Area Under the Curve (AUC) is computed, based on the probability estimates of the local lesion risk predictor

⁴ <http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> [accessed 10th of June 2018]

Table 2. Summary results LRS performance

Lesion Type	Patch Size	Mean AUC Region 1	Mean AUC Region 2
Emerging	10 x 4 x 10	0.7425	0.769
	20 x 4 x 20	0.7003	0.7144
	30 x 4 x 30	0.6739	0.6874
Changing	10 x 4 x 10	0.7607	0.7221
	20 x 4 x 20	0.7104	0.7491
	30 x 4 x 30	0.6976	0.7096

for the test patch using scikit learn⁵. We used thresholding to obtain a predicted label for visualisation and comparison. We have to point out that an exact matching of predicted label and subsequent annotation is not achievable, given a pre-stage and not required since we aim to predict risk of lesion growth or emergence and not to estimate exact segmentations of future lesions.

3.1 Evaluation Body Region 1

In Table 2 in column three the mean AUC for emerging and growing lesion types for body region 1 are summarized. For every lesion type three different patch sizes are evaluated. Figure 3 illustrates a prediction result for a growing lesion of a region 1 acquisition. The test image (left) is a transformed image from examination time point 003 to 004 using the longitudinal alignment approach introduced in Section 2. The extracted patch of this image in the region of the

⁵ http://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html [accessed 10th of June 2018]

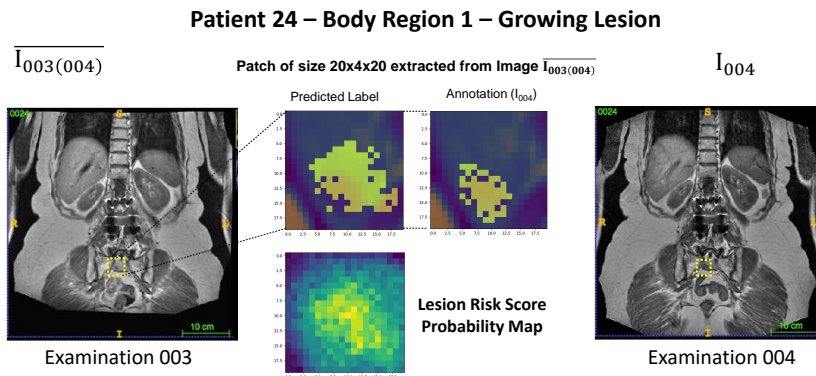


Fig. 3. Prediction of lesion growth from examination time point to time point in body region 1. The predicted label is visualised in the second column, below the underlying Local Lesion Risk Score probability map is shown and the manual annotation is visualised in the third column.

lesion visible in the image I_{003} is visualised in the first row in the center, with the predicted label in the second column and the annotation of the future lesion position extracted from image I_{004} in the third column. In the second row the predicted probability map of the local lesion risk score is visualised, where yellow shows regions of high probability and blue of low probability.

3.2 Evaluation Body Region 2

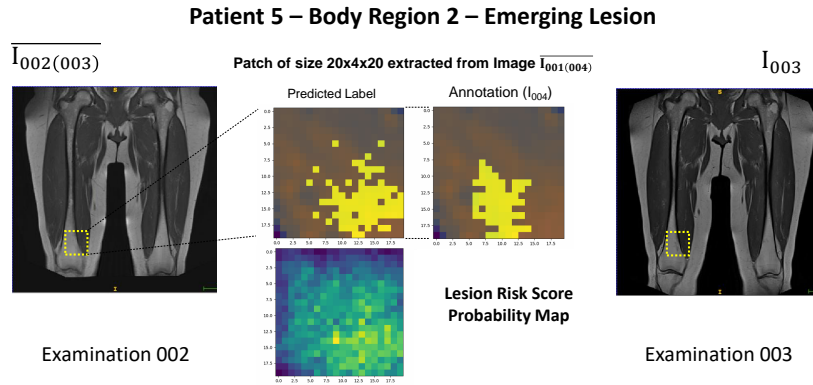


Fig. 4. Prediction of an emerging lesion from examination time point 002 to time point 003 in body region 2. The predicted label is visualised in column 2, below the underlying Local Lesion Risk Score probability map is shown and the manual annotation is visualised in column 3.

In Table 2 in column four the mean AUC for emerging and growing lesion types for body region 2 are summarized. For every lesion type three different patch sizes are evaluated. Figure 4 illustrates a prediction result for an emerging lesion of a region 2 acquisition. The test image (left) is a transformed image from examination time point 002 to 003. The extracted patch of this image in the region of the lesion visible in the target image I_{002} (right) is visualised, with the predicted label in the second column and the annotation of the future lesion position extracted from image I_{003} in the third column. In the second row the predicted probability map of the local lesion risk score is visualised, where yellow shows regions of high probability and blue of low probability.

3.3 Discussion

For both lesion types a decrease of the mean AUC is observable with increasing patch size, where emerging lesions show a steeper decrease as growing lesions in both bodyparts. Also growing lesions in the femur, knees or pelvis are better predicted than those in the thoracic or abdominal body region.

4 Conclusion

In this work we present a local Lesion Risk Predictor for accessing and visualising regions of high risk for bone lesions to emerge or to grow. We trained a random forest predictor using lesion image patches and annotations of subsequent lesions states of the longitudinal MR T1 weighted dataset observed. The main challenge here was to achieve an accurate longitudinal alignment of subsequent examination time points of a patient. To our knowledge this is the first attempt to predict bone infiltration patterns in MM using T1 weighted MR images. Current approaches focus on lesion detection (e.g. [11] for CT images) using deep learning techniques. We decided to investigate a lesion predictor based on a random forest classifier first, since its setup, parametrisation and evaluation is simpler compared to deep learning approaches. However, we incorporated the possibility to extend the proposed patch based approach to evaluate deep architectures for lesion prediction in the future. To this point prediction using the introduced local lesion risk score is limited to image patches. For future work we aim to adapt the proposed score to be able to predict probability maps for entire volumes. Additionally, we will incorporate different modalities, and data from additional bodyparts in the framework proposed to longitudinally model infiltration and also osseous destruction patterns caused by the progress of multiple myeloma.

Acknowledgement

This work was supported by the Austrian Science Fund (FWF) project number I2714-B31

References

1. M. A. Dimopoulos, J. Hillengass, S. Usmani, E. Zamagni, S. Lentzsch, F. E. Davies, N. Raje, O. Sezer, S. Zweegman, J. Shah, A. Badros, K. Shimizu, P. Moreau, C.S. Chim, J. J. Lahuerta, J. Hou, A. Jurczyszyn, H. Goldschmidt, P.r Sonneveld, A. Palumbo, H. Ludwig, M. Cavo, B. Barlogie, K. Anderson, G. D. Roodman, S. V. Rajkumar, B. G.M. Durie, and E. Terpos. Role of Magnetic Resonance Imaging in the Management of Patients With Multiple Myeloma: A Consensus Statement. *Journ. of Clin. Oncology*, 33(6):657–664, feb 2015.
2. B. G. M. Durie, R. A. Kyle, A. Belch, W. Bensinger, J. Blade, M. Boccadoro, J. Anthony Child, R. Comenzo, B. Djulbegovic, D. Fantl, G. Gahrton, J. Luc Harousseau, V. Hungria, D. Joshua, H. Ludwig, J. Mehta, A. Rodrique Morales, G. Morgan, A. Nouel, M. Oken, R. Powles, D. Roodman, J.s San Miguel, K. Shimizu, S. Singhal, B. Sirohi, P. Sonneveld, G. Tricot, B. Van Ness, and Scientific Advisors of the Intern. Myeloma Foundation. Myeloma management guidelines: a consensus report from the Scientific Advisors of the International Myeloma Foundation. *The Hematology Journal*, 4(6):379–398, 2003.
3. M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith. FSL. *NeuroImage*, 62(2):782–90, aug 2012.

4. J. K. Kloth, J.s Hillengass, K. Listl, K. Kilk, T. Hielscher, O. Landgren, S. Delorme, H. Goldschmidt, H.-U. Kauczor, and M.-A. Weber. Appearance of monoclonal plasma cell diseases in whole-body magnetic resonance imaging and correlation with parameters of disease activity. *Int. Journ. of Cancer*, 135(10):2380–2386, nov 2014.
5. R. A. Kyle and S. V. Rajkumar. Multiple myeloma. *Blood*, 111(6):2962–72, mar 2008.
6. L. Lambert, P. Ourednicek, Z. Meckova, G. Gavelli, J. Straub, and I. Spicka. Whole-body low-dose computed tomography in multiple myeloma staging: Superior diagnostic performance in the detection of bone lesions, vertebral compression fractures, rib fractures and extraskkeletal findings compared to radiography with similar radiation. *Oncology letters*, 13(4):2490–2494, apr 2017.
7. M.-V. Mateos, M.-T. Hernández, P. Giraldo, J. de la Rubia, F. de Arriba, L. L. Corral, L. Rosiñol, B. Paiva, L. Palomera, J. Bargay, A. Oriol, F. Prosper, J. López, J.-M. Arguiñano, N. Quintana, J.-L. García, J. Bladé, J.-J. Lahuerta, and J.-F. S. Miguel. Lenalidomide plus dexamethasone versus observation in patients with high-risk smouldering multiple myeloma (QuiRedex): long-term follow-up of a randomised, controlled, phase 3 trial. *The Lancet. Oncology*, 17(8):1127–1136, aug 2016.
8. M. Merz, T. Hielscher, B. Wagner, S. Sauer, S. Shah, M. S. Raab, A. Jauch, K. Neben, D. Hose, G. Egerer, M.-A. Weber, S. Delorme, H. Goldschmidt, and J. Hillengass. Predictive value of longitudinal whole-body magnetic resonance imaging in patients with smoldering multiple myeloma. *Leukemia*, 28(9):1902–1908, sep 2014.
9. M. Modat, D. M. Cash, P. Daga, G. P. Winston, J. S. Duncan, and S. Ourselin. Global image registration using a symmetric block-matching approach. *Journal of medical imaging (Bellingham, Wash.)*, 1(2):024003, jul 2014.
10. P. Tosi. Diagnosis and treatment of bone disease in multiple myeloma: spotlight on spinal involvement. *Scientifica*, page 104546, dec.
11. L. Xu, G. Tetteh, J. Lipkova, Y. Zhao, H. Li, P. Christ, M. Piraud, A. Buck, K. Shi, and B. H. Menze. Automated Whole-Body Bone Lesion Detection for Multiple Myeloma on 68 Ga-Pentixafor PET/CT Imaging Using Deep Learning Methods. *Contrast Media & Molecular Imaging*, 2018:1–11, jan 2018.