Computer-aided techniques for assessment of MRI-detected inflammation for early identification of inflammatory arthritis

Aizenberg, E.

6

Summary and general discussion

In this thesis, we developed several computer-aided methods for assessment of MRI-detected inflammation in patients with inflammatory arthritis. The described studies focused on the tasks of comparative visualization, automatic quantification, and feature selection, with the underlying aim of aiding early diagnosis of spondyloarthritis (SpA) and rheumatoid arthritis (RA).

Chapter 2 presented an interactive scoring tool for evaluation of inflammatory changes over time in patients with axial SpA. Locally-rigid image registration was applied to compensate for patient posture differences between scanning sessions and fuse baseline and follow-up MR scans into a single color-encoded image. The resulting visualization offered vivid distinction between areas of increase versus decrease in inflammation over time, coupled with automatic labeling of vertebral units (VUs) and an interactive scoring module whose entry fields were activated in synchronization with the VU selected by the reader in the image. Expert readers found that a key advantage of such computer-aided scoring was that it allowed for direct visualization and measurement of inflammatory changes from a single image, as opposed to two separate images. In addition, the synchronization between the image and the scoring module significantly decreased the chance of mistyping errors while filling out the digital scoring form. At the same time, the moderate inter-reader agreement on the degree of inflammatory change pointed to the need of further standardizing interpretation of such color-encoded visualization. To this end, automatic quantification of the degree of change could be the ultimate desirable goal.

Chapter 3 proposed a framework for automatic quantification of bone marrow edema on MRI of the wrist, for early detection of RA. To combine image data from the coronal and axial sequences into a single 3D image, super-resolution reconstruction was applied. The carpal bones were located using atlas-based segmentation and signal associated with bone marrow edema was identified by fuzzy clustering. Correlation between quantitative measurements and visual scores was assessed in a large cohort of early arthritis patients. The resulting measurements were largely consistent with visual scores, indicating that automatic quantification of bone marrow edema on MRI of the wrist is feasible. It was observed, however, that incomplete fat suppression during MRI acquisition can have an adverse effect on measurement accuracy, resulting in false detections. Solving this requires further improvement of the quantification method.

Chapter 4 extended and further developed the method of **Chapter 3** to measure tenosynovitis of the extensor and flexor tendons of the wrist. Atlas-based segmentation was used to locate the bones and place initial landmarks for tendon regions. These initial landmarks were then used as inputs for marker-based watershed segmentation. A measurement region of interest was defined around the tendons. As in **Chapter 3**, signal associated with inflammation was identified using fuzzy clustering, with the modification of obtaining a one-sided probability map. Correlation between quantitative measurements and visual scores was assessed in a large cohort of early arthritis patients. Strong correlation was observed, indicating that automatic quantification of tenosynovitis on MRI of the wrist is feasible. The study also brought out multiple challenges pertinent to the quantification task, such as moderate segmentation performance and sources of false detections. In particular, blood vessels and synovitis present with the measurement region of interest were two strongly contributing factors to a consistent offset in quantitative measurements.

Chapter 5 sought to identify whether the common set of 61 MRI-detected inflammatory features visually graded across the wrist, metacarpophalangeal, and metatarsophalangeal joints can be reduced to a smaller set of features specific for

RA, given the knowledge that some features are also frequently present in symptom-free persons. The difference in frequency of inflammation presence was studied between 199 RA patients and 193 controls. A subset of 30 RA-specific features (mainly locations of tenosynovitis and synovitis) was obtained by applying a cutoff on the frequency difference while maximizing discriminative performance. For validation, this subset was used to predict arthritis development in 225 clinically suspect arthralgia (CSA) patients. The smaller subset demonstrated comparable predictive accuracy to the original set. These results suggest that it is possible to preserve the diagnostic capacity of MRI with regard to prediction of progression from CSA to clinical arthritis while scoring only half of the features that are typically scored. In addition, this leads to new research questions about the processes driving the inflammation at the identified anatomical locations and whether this can help gain better understanding of arthritis pathogenesis.

General discussion

This thesis contributed towards computer-aided assessment of MRI-detected inflammation in patients with inflammatory arthritis. The feasibility studies of **Chapters 2–4** provide reference points for interactive comparative visualization in axial SpA and automatic quantification of inflammation in RA, while **Chapter 5** further elucidates the diagnostic role of individual inflammatory features in prediction of RA development. Results showcase the promise of computer-aided techniques to overcome the limitations of visual scoring discussed in **Chapter 1**. In particular, on the automatic quantification front, the observed correlations between quantitative measurements and visual scores are encouraging considering that the methods were validated in a large cohort of early arthritis patients. Nevertheless, these techniques are not yet sufficiently robust and precise to be used in clinical practice. We identified a number of key challenges that must be addressed on the path to achieving this goal.

First, improving segmentation accuracy is an important direction of future work, as it is an essential component of the presented image processing

frameworks. In comparative visualization of MRI of the spine, the accuracy of locally-rigid registration is directly dependent on the precision of vertebrae segmentation. Furthermore, the labeling of VUs and consequently the interactive scoring features are entirely dependent on whether all evaluated vertebrae were detected. Within the context of automatic quantification, segmentation accuracy affects the variability and reliability of quantitative measurements. In addition to overall accuracy improvements, mislabeling errors should be investigated in detail. To ensure wide applicability in clinical practice, segmentation and quantification should also be robust to variations in MRI acquisition protocols and scanners. One possible way to account for this, as part of an atlas-based framework, could be to form sub-atlases of images acquired under a range of echo/repetition times and magnetic field strengths. Then, prior to segmenting a target image, the most appropriate sub-atlas would be identified based on acquisition parameters recorded in the image's DICOM data. It should be noted that the underlying approach of segmenting by drawing on knowledge accrued in a set of annotated images that constitute an atlas, or more generally a training set, is not unique to atlas-based segmentation. The atlases of the methods presented in this thesis can be used with other knowledge-based techniques, such as active appearance models or convolutional neural networks.

The presented quantification methods for RA focused on bone marrow edema and tenosynovitis in the wrist joint. However, as indicated by the results of **Chapter 5**, features that are predictive for progression to clinical arthritis also include synovitis and are spread not only across the wrist joint, but also the metacarpophalangeal joints in the hand and metatarsophalangeal joints in the foot. Therefore, it is important to expand the quantification framework to these joints and include the measurement of synovitis. The atlas-based nature of the framework provides a straightforward path for including additional joints by adding manually annotated atlases of these joints to the wrist atlas. To incorporate synovitis measurements, inflammation could be measured in the anatomical regions bounded by bones and tendon regions, which are already segmented.

In this thesis, we assessed the consistency between quantitative measurements and semi-quantitative visual scores by evaluating the correlation between these measures. This allowed us to establish that automatic quantification of MRI-detected inflammatory features frequently seen in RA patients is feasible and is largely consistent with visual scoring. To go beyond feasibility and towards rigorous evaluation of true positive versus false positive detections of inflammation, it is important that future studies assess absolute agreement between regions identified as inflammation by quantitative measurements and ground truth manual segmentations of inflammatory features by human experts. This would demand a large investment of resources, since human experts would need to manually segment all voxels considered to be part of each inflammatory feature. However, such studies may be essential to demonstrate a convincing level of agreement between automatic techniques and human experts, in order to facilitate the use of such computer-aided methods in clinical practice.

Ideally, application of automatic methods as part of a future clinical routine should be possible directly after acquisition of an MR scan, as soon as the DICOM image is stored in the patient database. However, at the present this would not be possible due to a number of artifacts that occur during acquisition and require correction prior to running quantitative analysis. For example, form entry errors can be made in the DICOM fields with regard to which location was scanned (e.g. wrist, foot) and on which side (left, right). A mistake in one of these fields will cause the atlas-based frameworks to use the wrong atlas for segmentation and result in failed quantification. Another issue is that MR scans acquired with a frequency-selective fat-saturated sequence sometimes suffer from fat suppression inhomogeneity, which may cause quantitative measurements to confuse regions of fat tissue for inflammation. There can be two general approaches to addressing these acquisition issues: 1) improve acquired image quality requirements and systematically minimize the possibility of DICOM field errors through stricter protocols and software interface, or 2) develop automatic methods for handling acquisition issues as part of the overall computer-aided framework. The choice

boils down to resource management. As computer-aided techniques become more integrated in clinical practice, a balanced approach could be to let operating technicians manage those acquisition issues that can be fixed with minimal additional inconvenience and cost for patients at acquisition time. On the other hand, correction of acquisition issues that would substantially increase patient inconvenience, cost, and procedure time can be delegated to the image processing engineers.

In conclusion, this thesis has explored the prospect of computer-aided assessment of MRI-detected inflammation for early identification of inflammatory arthritis. The presented studies showcase the potential of comparative visualization and automatic quantification to overcome the limitations of visual scoring and lay out a fertile ground for future improvements. Additionally, the understanding of the diagnostic role of individual inflammatory features in prediction of RA development is further advanced. Collectively, these findings can help facilitate the use of MRI for early diagnosis of inflammatory arthritis and potentially increase chances of better outcome and quality of life for patients.