Analysis of knee osteoarthritis data via bioinformatics tools

1,2 Margarita Kotti, 1 Lynsey D. Duffell, 2,3,4 Aldo A. Faisal, and 1 Alison H. McGregor

1Musculoskeletal (MSK) Laboratory, Division of Surgery, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, Charing Cross Hospital, London W6 8RF, UK
2 Brain Behaviour Laboratory, Department of Bioengineering, Imperial College London, SW7 2AZ London, UK
3Department of Computing, Imperial College London, SW7 2AZ London, UK
4MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, London, UK

Knee osteoarthritis is the second leading cause of disability, with rates rising as a consequence of the expanding ageing population. With a view to maintaining an active and independent older population our lab is exploring the analysis of gait patterns across a spectrum of osteoarthritic patients to identify early biomarkers of the disease progress. Currently, we have a large volume of raw sensor data captured via a motion capture system, force plates, and electromyography. These high-dimensional, multimodal biomedical datasets are being processed using bioinformatics approaches to reveal the implicit information they contain.

More specifically machine learning techniques have been exploited to (i) detect gait patterns produced by osteoarthritis subjects and (ii) extract parameters capable of differentiating between normal and osteoarthritic knees. Initial experiments exploited the joint angles at the pelvis, hip, knee, and ankle, as well as the direction of the foot during walking. The aforementioned parameters of 16 subjects were analysed by a support vector regressor. Two subjects were identified as demonstrating a gait pattern closer to the osteoarthritic model despite identifying themselves non-osteoarthritic. Currently, we are exploiting a larger dataset comprising the ground reaction forces of 94 subjects. For this dataset, random forests were utilised in order to detect discriminating parameters with physical meaning and to induct rules. The bioinformatics tool evolved was able to identify approximately 20% of our “normal” population as having a high risk of knee osteoarthritis. The relevance of this in relation to a non-invasive early marker of disease requires further exploration.