OATECH meeting 17th December 2017 summary of imaging discussion

Should we cover all imaging modalities?

It was agreed all types of imaging methods should be considered but there might be different inclusion criteria depending on imaging modality. In particular some advanced techniques (e.g. SPECT/CT) may have smaller patient numbers but have significant implications for the directions for future studies, or a better assessment of the response of DMARDs, or suggest where there is funding "hole". To specifically exclude some would also be a bias unless we have preliminary hypothesis.

When including all data we should include assessments of the reliability, sources of variability, and how well macroscopic imaging measures relate to the underlying microscopic pathological changes, do we need better methods not yet available or in development? Can we make any recommendations?

Since papers are continuously being published and methods are under continual development any "snapshot" review is out of data as published. Can AI be used to actively update a database with appropriate papers and provide a dynamic picture of research methodology and results?

A) How shall we stratify the modalities in terms of usefulness?

Difficult to exclude modalities without risk of bias and there is also a methodological shift in accuracy, reliability and capability of imaging technology. It may be useful to attempt definition of an information content of different imaging methods. The results of the final analysis may give an indication of "usefulness" – what wasn't discussed was aspects of ease of use, cheapness, cost-benefit, ready applicability to large cohort studies.

B) What questions can we ask of the database re imaging:

- correlation w.r.t. specific scores
- correlation to outcome
- sensitivity to progressive change

Sensitivity to progressive change would be important since OA is a progressive disease. Overall thought that any parameter that can be measured is something that could be used, but see answer to C.

C) What criteria for including imaging results

Important factors include consistency of the measures, particularly in the light of technological advances that occur over a decade or two, but data from 20 yrs may be quite a primary result to include, as well as large early trials. QC assessment of individual papers would likely best be done by a minimum dataset that is reported that is the threshold above which papers are included. Parameters such as: well defined and detailed imaging protocol; machine type; number of readers; scoring system details. Quantitative measures were argued by some as the most relevant, but partial volume effects at poor resolution hamper these. Highly detailed scoring may simply add noise as opposed to more gross scoring of grades 0 to 4 for example. This is an aspect that needs to be explored in a subset of papers to develop our inclusion criteria thresholds.

D) Should we do meta-analysis w.r.t. features (e.g. BML)

- how will we combine results
- Quantitative or semi-quantitative
- Cross-sectional or longitudinal

Meta-analysis when comparing the same score, but with the caveat that we believe over the period of the studies included there are not technological or methodological changes that would invalidate the comparisons. Important to understand the source of noise and whether there are biases in any parameter being studies. This could be important in longitudinal studies for example if readers are not blinded to the sequence of the data.

Combining different data would be problematic. However the results of the analysis may suggest how data from different modalities would fit together, as part of a decision tree for example, particularly with reference to what other imaging data (e.g xray, US) was used to provide inclusion/exclusion to a subsequent imaging study (e.g. MRI or and advanced technique).

SUMMARY POINTS

Since all types of imaging will be considered there need to be strict inclusion criteria with respect to a variety of factors including: consistency of data and measurements; quality of data; details of methodology and analysis; equivalence of data.

Since some of the more recent methods, or more complex/expensive methods, may be more limited in patient study numbers used there will likely be two main outcomes:

- Main meta-analysis that uses studies with minimum N= 100 and likely includes only US, MRI and X-ray methods. The focus will be on the most widely used imaging features that are used and related to clinical and outcome measures. (e.g. BMLS, effusion, synovitis, JSW, osteophytes, cartilage damage). Recommendations could include how to improve imaging protocols, better understanding of the strengths/weaknesses of current methods, agreements on what are the best parameters to measures and how to measure them.
- Advanced technology analysis that will likely include SPECT/CT, PET, some quantitative MRI (eg. T1rho); image analysis with ML methods such as texture and with multi-parametric analysis; a discussion of how well/or not imaging relates to actual pathology – so may included ex vivo and non-human studies. Such an analysis could also lead to recommendations of the way forwards, funding needs, identifying “holes”.

<u>Clinical Trials Session</u>

In this session, delegates were asked to consider two key questions to inform the group’s systematic review. Firstly they were asked to consider definitions of osteoarthritis. This would provide an informed benchmark to the eligibility of osteoarthritis trials in the systematic review. Whilst there was little direct reference to international guidance on osteoarthritis definitions using guidance from organisations such as EULAR or OARSI or ACR, there was an overwhelming association that definition should be based on a combination of clinical features. Radiological definition was acknowledged to be important for older studies, but it was clear that more contemporary evidence may place less

emphasis on this criterion. Some evidence of joint pathology by biomarker of some description e.g. imagining was desirable but the group also felt that being able to exclude other diagnoses, particularly from hip osteoarthritis, would be advantageous.

There was an important suggestion that risk factors, in this analysis, could be analysed collectively for the development of osteoarthritis but also by phenotype, classifying participants by their predominant feature. Through this, it would be possible to better understand the potential risk factors for subgroups of a heterogeneous population.

The second question posed to delegates was to ask them to identify candidate risk factors for the development of pain (two groups) and structural change (two groups) for people with osteoarthritis. This would provide a working list of candidate factors which could be cross-referenced to the systematic review as we validate the results from the evidence. The potential candidate factors are listed below:

Risk factors for pain

- Psychological – including catastrophisation
- Environmental/social – including cultural/occupational/weather
- Physiological – including BMI and pain centralisation
- Structural – including bone marrow lesions
- Co-morbidities – such as diabetes and heart disease

Risk factors for structural change

- Obesity
- Joint malalignment
- Gender
- Musculoskeletal function – muscle control/ligament integrity/innervation
- Previous injury – particularly to meniscus and ACL
- Comorbidities
- Occupational and social factors

<u>Pain and Function session</u>

In this session, delegates were asked to consider two key questions to develop the team's systematic review. Firstly, they were asked to discuss if only clinically meaningful measures of pain and function that are widely used in studies e.g. Numerical Rating Scale (NRS) for pain, Visual Analog Scale (VAS) or WOMAC, EuroQol, are needed in our searches of hip and knee OA studies. While there was a recognition by the group that many studies include validated pain scores as primary or secondary outcome measures in studies, in addition to measures of function e.g. EuroQol, WOMAC function and stiffness subscales, the group considered that there are wider factors contributing to the pain and functional impairment in OA.

There was an important suggestion that WOMAC pain, stiffness and function are strongly correlated to each other, and recognising that distinct factors that contribute to pain and impaired function could also be assessed by the review. These feed into the previous discussion in the 'Clinical Studies' session of risk factors for pain.

The second question considered which additional measures the group considered to be important for measuring pain and function in OA. The group identified additional pain and function measures which are summarised below:

Additional measures for pain

Quantitative sensory testing including temperature, pain pressure thresholds, temporal summation

Gait analysis

Time to joint replacement from first symptom onset

Use of analgesic medication including but not limited to NSAIDs, opiates and centrally-acting analgesics

Functional brain neuroimaging to understand pain pathways/networks in OA

Use of more specialised questionnaires e.g. painDETECT to obtain measures of neuropathic/inflammatory pain

Additional measures of function

Walk time e.g. 5 minute walk test

Stair climb test

Data on mobility using FITBIT /Smartphone type devices

Cost effectiveness of joint surgery e.g. use of healthcare services after surgery

Need for other joint replacements e.g. contralateral knee, hip etc

In the last question of the session, delegates were asked if they think it is important to combine pain and functional measures in the systematic review, to which the overwhelming response was 'yes'.