

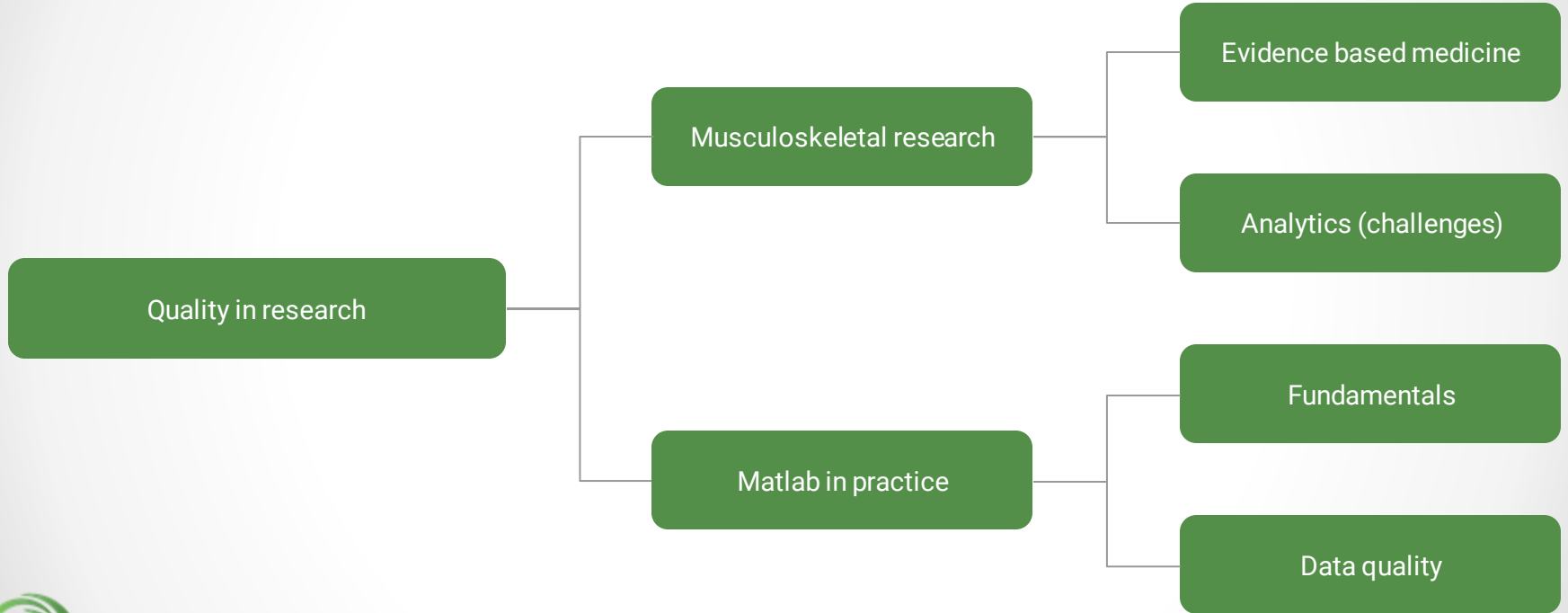
# Crawling before running: Advanced analytics in orthopaedics research with Matlab

Focus on data quality

Corey Scholes, PhD. (Lead Consultant)  
Milad Ebrahimi Beng (Hons) (Lead Engineer)



# Outline



# Orthopaedic and musculoskeletal conditions

- Encompasses trauma, acute injury and chronic disease
  - Road traffic accident
  - Workplace injury
  - Sports injury
  - Osteoarthritis, tendinopathy, osteopenia
- Nearly every person affected by a musculoskeletal condition at least once in their life
- Social cost of osteoarthritis alone 0.25 - 0.5% of GDP ([Puig-Junoy et al 2015](#))
  - ~\$60billion/annum in Australia



# Orthopaedic treatments

- Effective diagnosis of injury or chronic diseases
- Spectrum of treatment options from non-invasive to very invasive
- Considerations of patient preference, cost, clinical benefit and nature of the condition
- Often applied in combination
  - Ambiguity in best treatment for a specific patient



Soft tissue repair



Exercise therapy



Stem cell injection



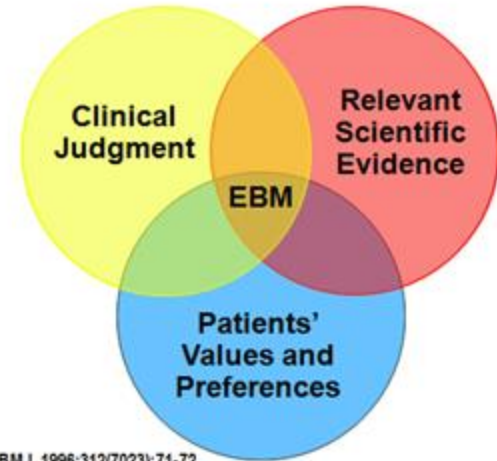
Total joint replacement



# Evidence based medicine

- Attempts to replace folklore, tradition and theoretical reasoning
- Scientific evidence
  - Trials
  - Studies
  - Analyses
  - Systematic reviews
  - Meta-analyses
  - Consensus and practice guidelines
- Patient-centred
  - Expectations
  - Self-reported outcomes

## What Is Evidence-Based Medicine?



Sackett DL, et al. BMJ. 1996;312(7023):71-72.



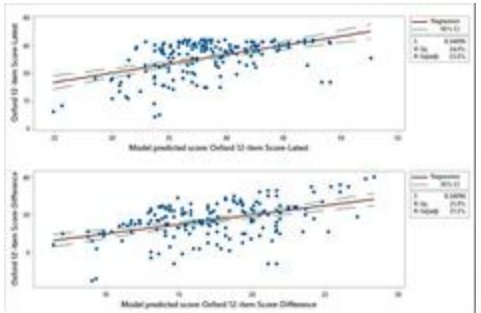
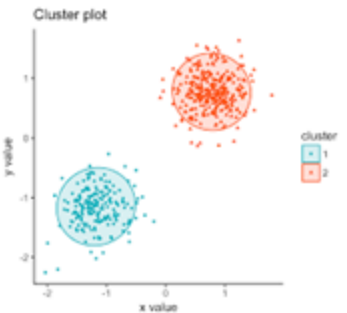
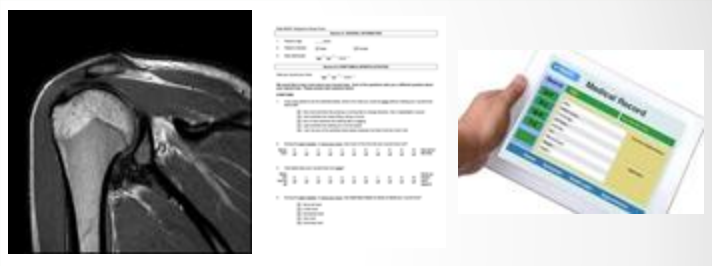
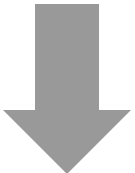
# Our role



# Tools and applications



Patient registries



# Analytics

- Underpins evidence generation in EBM
- Required to inform shared-decision making between clinician and patient
- Analytics alone  $\neq$  insight or practice change
- Integration within clinical practice



$$\text{Value} = \frac{\text{Patient outcome}}{\text{Healthcare cost}}$$

**Figure 1.** A continuous learning healthcare system.





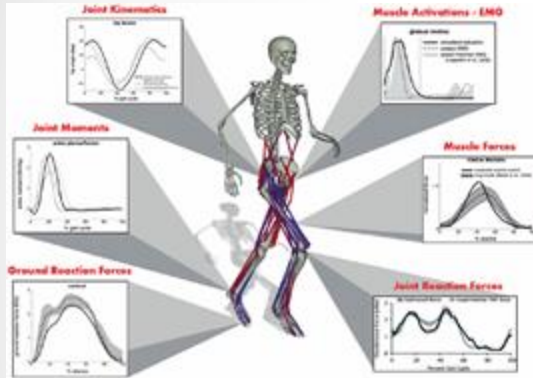
# Challenges

...

Focus on quality



# Biomedical research



[Ku et al 2015](#)

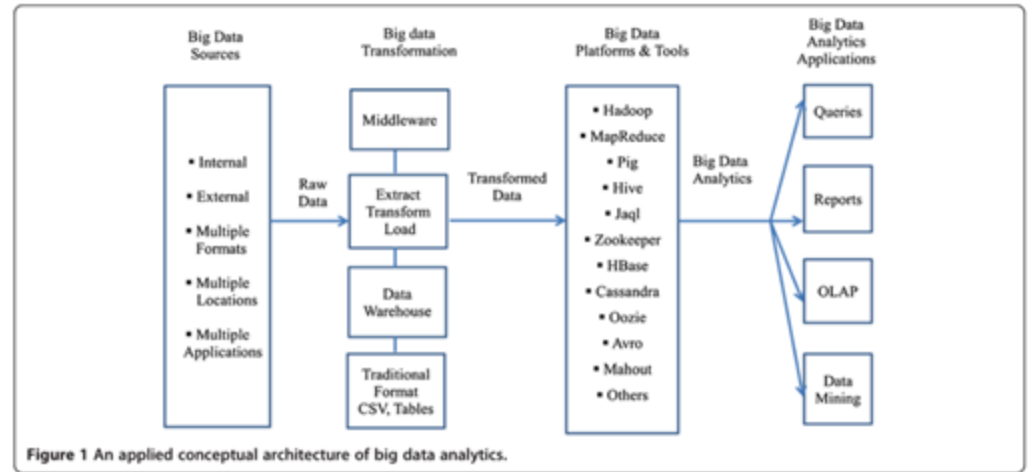


- Generate relevant scientific evidence for use in clinical practice
- Total global investment of \$240billion in health related research and development ([Rottingen et al 2015](#))
- Most clinical research is not useful
  - Replication
  - Insufficient design
  - Inadequate planning
  - Biased methods/interpretation



# Analytics and big data

- Petabytes of health data collected daily by myriad of health providers and associated organisations
  - Government departments
  - Hospital
  - Insurance
  - Device industry
  - Private practice providers
  - Not-for-profit organisations
- 4 V's ([Raghupathi et al 2014](#))
  - Volume
  - Velocity
  - Variety
  - **Veracity**



[Raghupathi et al 2014](#)



# Data quality

## Fundamentals

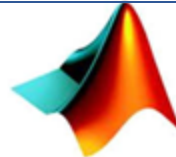
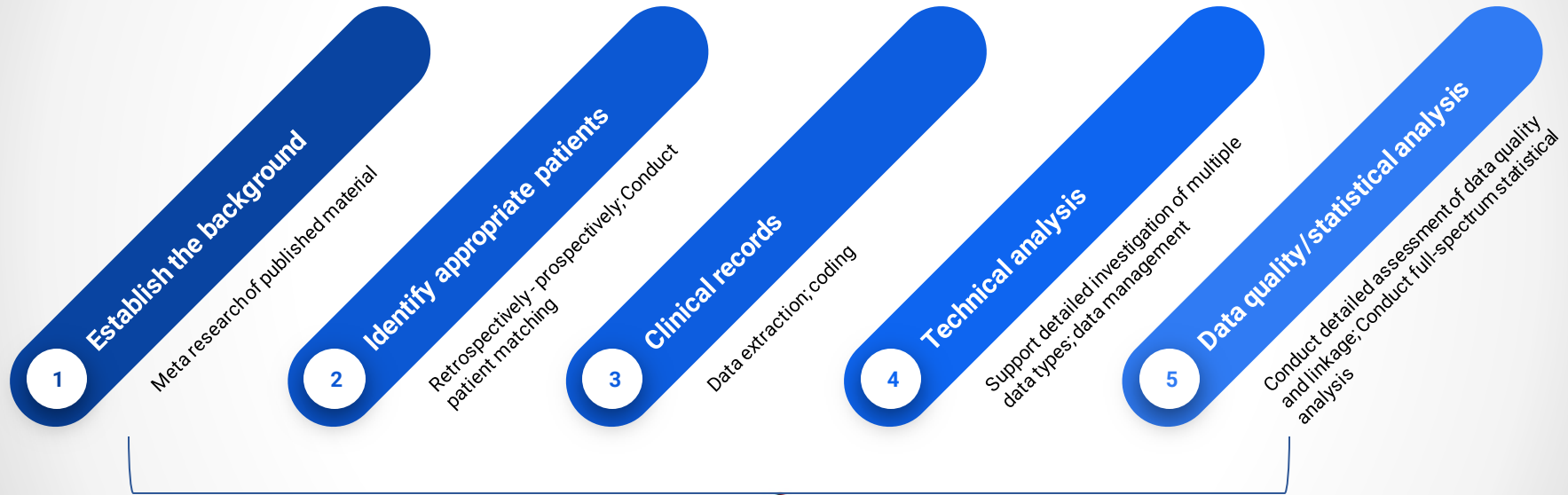
## Logistics

1. Do we have *all* the relevant patients for analysis?
2. Have they been classified correctly?
3. Do we have complete and accurate data?

1. Interoperability of software
2. Data in silos
  - a. Definitions
  - b. Culture
3. Time and labour requirement to build datasets



# Research process



# Matlab in Practice

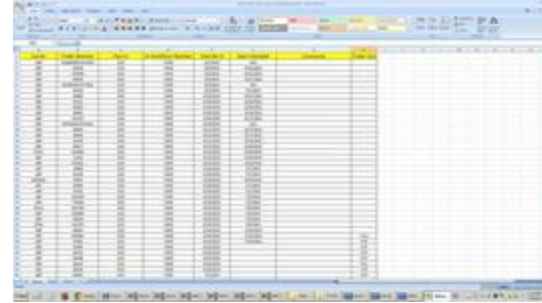
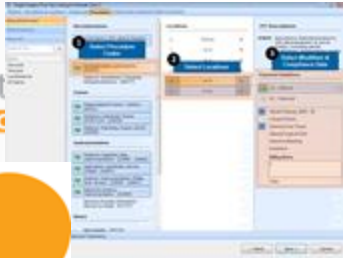
...

Research applications

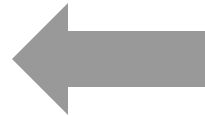


# Do we have the right patients?

Patient management systems



Research databases/systems



REDCap  
Research Electronic Data Capture

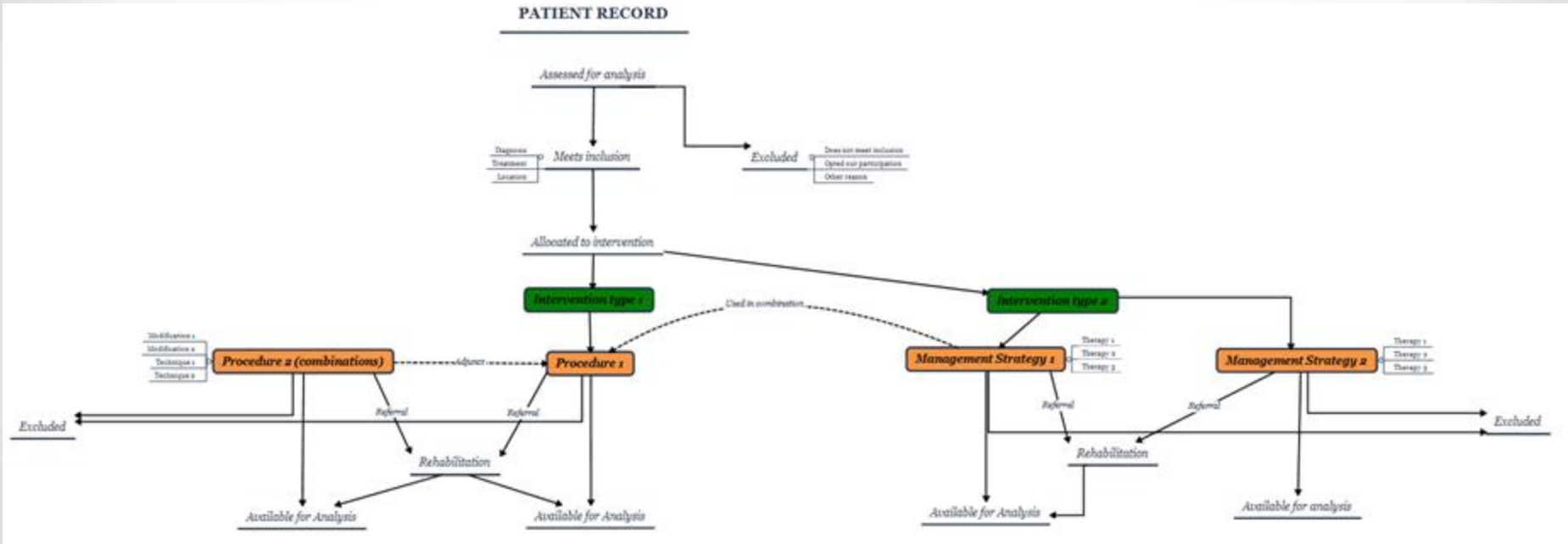


berd™



EBM ANALYTICS

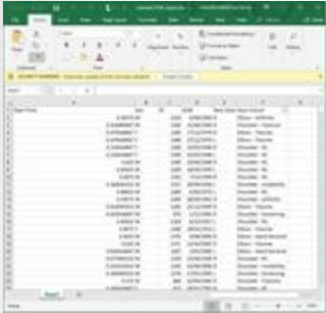
# Have we classified patients correctly for analysis?



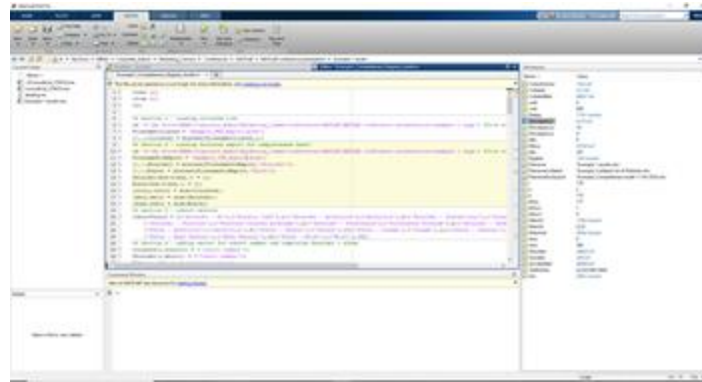
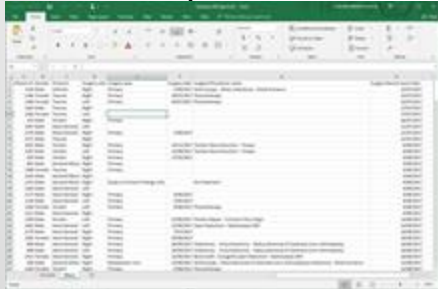


# Example 1

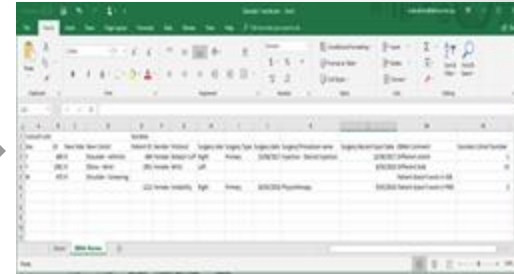
PMS spreadsheet



rDB spreadsheet



Output spreadsheet

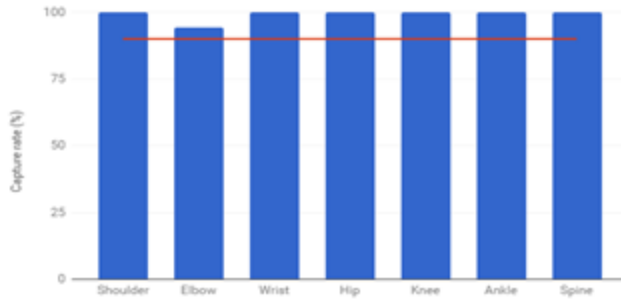


Results application

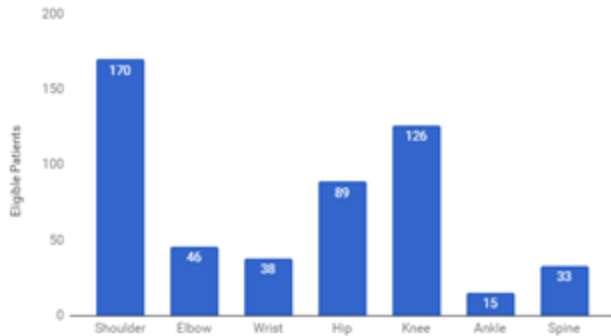


# Results and output

Registry capture rate (%) of eligible patients during audit period



Eligible patients (N) during audit period



## Application

- Regular reporting to stakeholders
- Actionable information to iterate processes
- Constant communication with
  - Collecting
  - Interpreting
  - Decision-makers



# Discussion

## Key Points

- Source of errors
  - Patient compliance
  - Staff compliance
  - Clinical situation
  - Human error
- Lack of automation
- Process evolution

## Key Lessons

- Define inclusion/exclusion criteria
- Regular logic checks
  - Check each patient in the group
- Establish fast feedback with regular reporting



# Musculoskeletal imaging

## Magnetic resonance imaging



## Key points

- Non-invasive method for visualization of anatomy
- Stimulates water molecules in tissues using strong magnetic fields
- Modification of key parameters alter contrast, sharpness and quantitative information

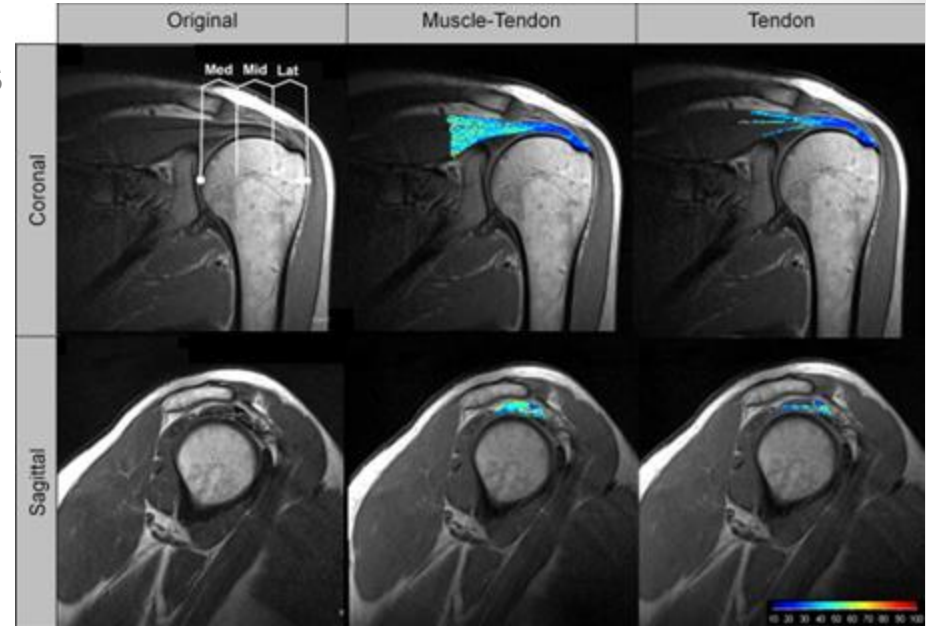


# Do we have accurate data for analysis?

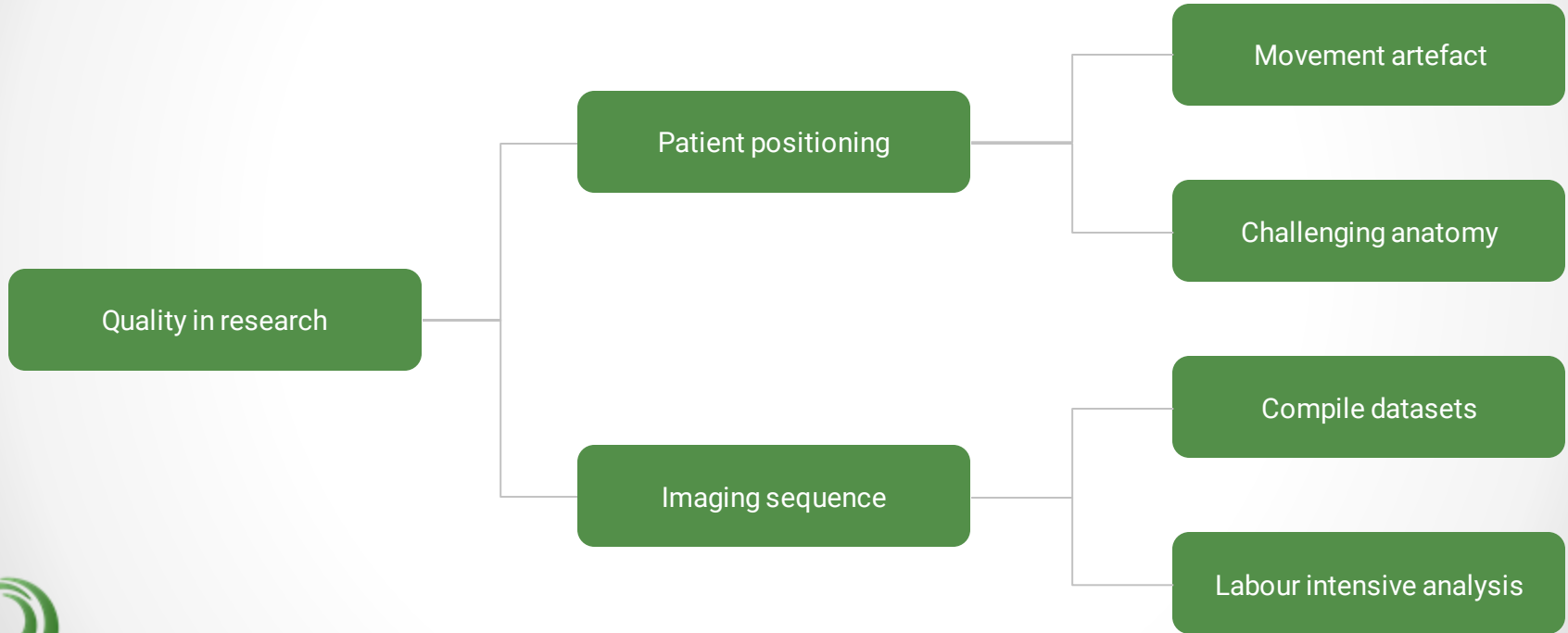
## Imaging analysis

- Musculoskeletal medicine utilises diverse data
- Imaging data is large, complex and vulnerable to poor quality
- Quality issues can be invisible until too late
  - Analysis at scale

## Quantitative mapping



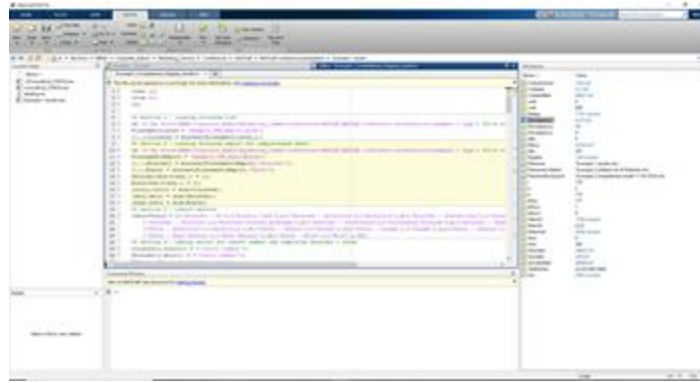
# Quantitative imaging



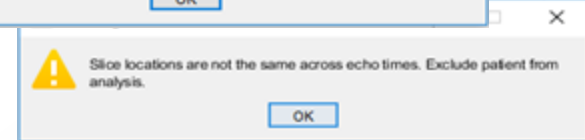
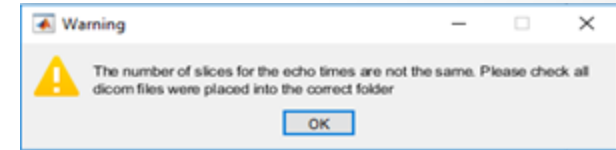
# Example 2



Dicom metadata



Output



# Discussion

## Key Points

- Source of errors
  - Patient movement
  - Software error
  - Machine malfunction
  - Calibration error
- Lack of automation
- Scale vs quality

## Key Lessons

- Clear understanding of sequence and analysis requirements
- Communication with radiography personnel
- Establish solid understanding of latest advances/literature





# Conclusions and Future

...

Challenges and applications



# Conclusions

## Problems solved

- Musculoskeletal medicine and research is data rich
  - Lacking quality
- Labour intensive tasks a barrier to insights and actionable information
- Increased transparency
  - Automation

## Key Lessons

- The ability to apply 'big data' techniques remains limited
- Poor quality is contributing to imprecise findings within the literature
- More work is required to improve data quality across a range of areas



# Future directions

## Challenges

- Countering misleading narratives
  - Poor methodology
  - Lack of quality control
- Disconnect between available treatments/technology and knowledge base
- Time taken to produce quality evidence in EBM

## Applications and use

- Distributed deployment of tools
- Broader access to contributors and stakeholders
  - Cloud-based infrastructure
- Refine quality control and pair with emerging analytical tools



# Acknowledgements

- Staff and students
- Clinical partners
- Patients and participants
- Mathworks

