A Synovial Fluid Fingerprint for End-stage Knee Osteoarthritis



Chethan Jayadev

Wadham College

University of Oxford

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Trinity 2014

For Anita, Rohan & Nikhil

“Research is to see what everybody else has seen,
and to think what nobody else has thought”

- Albert Szent-Györgyi

Acknowledgments

This has been a long and arduous journey. Fortunately, it is not a journey I had to make alone.

First and foremost, my deepest gratitude goes to my supervisors, Andrew Price and Philippa Hulley, for their unwavering support and guidance. Their friendship, encouragement and ability to positively channel my foibles have steered me towards completing this DPhil. Andrew has been a wonderful mentor to whom I can turn for advice regarding research, orthopaedics and life in general. Philippa has been able to successfully nurture orthopedic surgeons in the laboratory. Her intelligence, calm enthusiasm, patience and razor sharp wit have been truly inspirational. I must also sincerely thank Sarah Snelling for her advice and support at the lab bench.

When embarking on this DPhil, I was naive to the considerable expense of basic science research. I am therefore extremely grateful to the Oxford NIHR Biomedical Research Unit, Orthopaedic Research UK and the League of Friends at the Nuffield Orthopaedic Centre for funding my research.

One of the pivotal moments of this research was the collaboration with GlaxoSmithKline facilitated by Peter Taylor, to whom I am eternally indebted. I would especially like to thank Katherine Sully and Fiona Germaschewski at the GSK Biomarker Group. I have also received a great deal of generosity and technical advice from industry during my research. I am very grateful to Julie Wood (Life Technologies), Chris Walker (MesoScale Discovery), David Walker (Bio-Rad) and Florence Storey (R&D Systems). I am also thankful to Gary Collins for checking my statistical approaches and reassuring me I hadn't got it all wrong.

The Botnar Research Centre is a unique community of scientists and clinicians, where there was a healthy balance between life and research. My time there has rewarded me with many friends and colleagues. In particular Raj Rout, Luke Jones and Abtin Alvand have made my time in research immensely entertaining as well as intellectually stimulating. There are several others I must thank for their time and friendship: Ben Davies, Cameron Brown, George Grammatopoulos, Ben Kendrick, Nicholas Bottomley, Geraint Thomas, Richard Murphy, Stephen Mellon, Monica Armengol, Annie Ng, Catherine Swales and David Mahoney to name but a few. I must apologise to anyone I have omitted. Sandra Regan, Jo Brown, Barbara Marks and Jenny Ferris are the unsung heroes of the Botnar Research Centre who not only provided administrative support but also pastural care over coffee, biscuits and cake.

Finally, I thank my family. My mother and father have given their only son every opportunity I have ever wanted and often more than I have deserved. They have encouraged and supported all my endeavours. I am also sincerely thankful to my parents-in-law who have gone out of their way to help my young family, especially when I needed time and space to write this thesis.

This thesis is dedicated to my wife, Anita, and our two amazing little boys, Rohan and Nikhil. Anita and I married just a few months before starting my DPhil. Rohan arrived about half-way through my research time at Oxford and Nikhil entered the scene during my write-up while back in London. The DPhil has therefore been a constant and often disruptive feature of our lives together. Anita has been my rock and has supported me through some tough times. She has single-handedly raised our two sons and organised our lives, while I have been raising this DPhil. Her love, encouragement, selflessness and patience are without limits. The boys have been my light and inspiration during some dark times and they remind me what is important in life.

Abstract

A Synovial Fluid Fingerprint for End-stage Knee Osteoarthritis

Chethan Jayadev — Doctor of Philosophy
Wadham College — Trinity 2014

Osteoarthritis (OA) is a common and debilitating condition with no cure. The lack of disease-modifying treatments is linked to a deficiency of suitable biomarkers. This thesis focuses on synovial fluid (SF) as a source of combination biomarkers in knee OA. The thesis aims to demonstrate the utility of wide-spectrum SF analysis by defining a diagnostic "fingerprint" for end-stage knee OA from a diverse range of soluble molecules with pathophysiological relevance.

SF is a difficult matrix to study and volumes from OA joints are often insufficient for multiple marker analysis by conventional assays. This thesis begins by optimising, validating and comprehensively comparing SF analysis on two leading multiplex platforms, Luminex xMAP and MesoScale Discovery (MSD). Hyaluronidase treatment of SF permitted better intra-assay precision and higher assay signals compared to neat or diluted samples on the Luminex platform, but not conclusively on the MSD platform. Both platforms demonstrated acceptable precision and accuracy, but important differences were highlighted. Limited agreement in concentration measurements across platforms suggests that biomarker analysis is at best only quasi-quantitative.

Important practical considerations of SF sampling were addressed using multiplex analysis of 22 different markers in end-stage knee OA. The anatomical site of sampling was shown not to affect assay measurements even when there were significant compartmental differences in structural burden of disease. A longitudinal study demonstrated strong to excellent agreement in analyte measurements taken $\sim$2 weeks apart, during which there had been no change in clinical and radiographic severity. These studies suggest biological equilibrium in SF composition between knee joint compartments and short-term temporal stability in the SF biological profile in end-stage knee OA.

Building on these analytical and practical considerations, the SF biology of end-stage knee OA was explored with a view to biomarker discovery. Wide-spectrum SF analysis of a diverse range of 35 cytokines, chemokines, growth factors, matrix enzymes and markers of matrix turnover was conducted in 98 well-characterised patients with end-stage knee OA, knee injury and inflammatory arthritis. The data was analysed using both classical and multivariate statistical approaches.

Surprisingly, patients with cumulatively more severe structural phenotypes of knee OA demonstrated no differences in the concentration or variability of markers despite significant differences in structural burden of disease. This suggests end-stage knee OA has a constant biology and may be considered a useful terminal benchmark for OA biomarker research.

Partial least squares discriminant analysis produced detailed and streamlined multivariate models with excellent (>95%) diagnostic accuracy, reliability and specificity for patients with end-stage knee OA, knee injury and inflammatory arthritis. The models were successfully validated with 10 new end-stage knee OA test samples. The streamlined 8-marker model included a combination biomarker fingerprint for end-stage knee OA comprising "ridges" of PIIANP, TIMP-1, ADAMTS-4 and MCP-1 with "troughs" of IP-10 and TGFβ-1. These findings represent a new breed of biomarker that could qualify as a "virtual knee replacement". The results and multivariate approach used in the thesis serve as a proof-of-concept to stimulate a new era in OA biomarker discovery.

Contents

1 General Introduction **1**

1.1 The Synovial Joint 2

1.1.1 The Knee 2

1.1.2 Articular Cartilage 3

1.1.3 Subchondral Bone 13

1.1.4 Synovium 14

1.1.5 Synovial Fluid 14

1.1.6 Meniscus and Other Structures 15

1.2 Osteoarthritis 17

1.2.1 Epidemiology 17

1.2.2 Definition 17

1.2.3 Aetiology 18

1.2.4 A New Paradigm 20

1.2.5 Pathophysiology 20

1.2.6 Clinical Features and Diagnosis 30

1.2.7 Assessment of Symptoms 31

1.2.8 Assessment of Structure 32

1.2.9 Symptom - Structure Discord 36

1.2.10 Phenotypes 37

1.2.11 A New Staging Paradigm 38

1.2.12 Treatment 39

1.2.13 Lack of Disease-Modifying Therapies 40
1.3 Biomarkers 42
1.3.1 Need for Osteoarthritis Biomarkers 43
1.3.2 Classification and Terminology 43
1.3.3 Biomarkers Sources 46
1.3.4 Wet Biomarker Kinetics 47
1.3.5 Which Biological Fluid? 47
1.3.6 Observations on Current Trends 50
1.3.7 Whither Osteoarthritis Biomarkers? 51
1.3.8 The Road Forward 54
1.4 Summary 57
1.5 Scope of the Thesis 58
1.6 Detailed Objectives 58
1.7 Thesis Structure 59

2 General Methodology **62**
2.1 Introduction 62
2.2 Ethics 62
2.3 Patient Cohorts 63
2.4 Exclusion Criteria 65
2.5 Patient Assessments 65
2.5.1 Clinical Evaluation 66
2.5.2 Radiological Evaluation 66
2.5.3 Intra-operative Assessments 68
2.6 Synovial Fluid 68
2.6.1 Sample Collection and Storage 68
2.6.2 Preparation for Analysis 69
2.7 Immunoassays 69
2.8 Quality Control and Assay Performance 70
2.8.1 Standard Curves 70
2.8.2 Assay Performance Parameters 70
2.9 Statistics and Software 71

I Methods Development and Practical Considerations 72

3 Synovial Fluid Treatment To Improve Immunoassay Performance 73

3.1 Introduction . . . 73

3.2 Null Hypotheses . . . 74

3.3 Methodology . . . 75

3.3.1 Synovial Fluid Treatments . . . 75

3.3.2 Performance Parameters . . . 75

3.3.3 Statistical Methods . . . 77

3.4 Results . . . 79

3.4.1 Luminex Assays . . . 79

3.4.2 MesoScale Discovery Assays . . . 82

3.5 Discussion . . . 86

3.6 Conclusions . . . 91

4 Multiplex Analysis of Synovial Fluid: A Comparison of Platforms 92

4.1 Introduction . . . 92

4.2 Null Hypotheses . . . 93

4.3 Methodology . . . 94

4.3.1 Immunoassays . . . 94

4.3.2 Sample Spiking . . . 94

4.3.3 Assay Performance Parameters . . . 97

4.3.4 Cross-Platform Agreement . . . 98

4.4 Results . . . 100

4.4.1 Luminex Bead Events . . . 100

4.4.2 Limits of Detection . . . 100

4.4.3 Limit of Quantification . . . 103

4.4.4 Spike Recovery . . . 108

4.4.5 Intra-assay Precision . . . 110

4.4.6 ELISA Validation 110
4.4.7 Cross-platform Agreement 112
4.5 Discussion 126
4.5.1 Luminex Bead Events 126
4.5.2 Precision 127
4.5.3 Limits of Detection and Quantification 127
4.5.4 Spike Recovery and Accuracy 130
4.5.5 ELISA Validation and Cross-platform Agreement 131
4.6 Conclusions 135

5 Anatomical Site of Synovial Fluid Sampling 137

5.1 Introduction 137
5.2 Null Hypotheses 138
5.3 Materials and Methods 139
5.3.1 Site-specific Synovial Fluid Collection 139
5.3.2 Radiographic Assessment 139
5.3.3 Synovial Fluid Analysis 139
5.3.4 Internal Validation 140
5.3.5 Statistical Analysis 140
5.4 Results 142
5.4.1 Radiographic Assessment 142
5.4.2 Quality Control of SF Assays 144
5.4.3 Internal Validation 144
5.4.4 Reproducibility of Compartment SF Analysis 145
5.5 Discussion 153
5.6 Conclusions 156

6 Short-term Temporal Stability of Synovial Fluid Analysis 157

6.1 Introduction . . . 157

6.2 Null Hypothesis . . . 158

6.3 Methodology . . . 159

6.3.1 Ethics and Patient Recruitment . . . 159

6.3.2 Sample Collection and Patient Assessments . . . 159

6.3.3 Reference Group (External Validity) . . . 160

6.3.4 Synovial Fluid Analysis . . . 160

6.3.5 Statistical Analysis . . . 161

6.4 Results . . . 165

6.4.1 Clinical Assessments and External Validity . . . 165

6.4.2 Quality Control of SF Assays . . . 165

6.4.3 Repeatability of Interval SF Measurements . . . 168

6.5 Discussion . . . 177

6.6 Conclusions . . . 181

II Wide-spectrum Synovial Fluid Analysis 182

7 Part II Study Design and Methods 183

7.1 Introduction . . . 183

7.2 Patient Cohorts . . . 183

7.3 Patient Assessments . . . 184

7.4 Synovial Fluid Preparation . . . 185

7.5 Biological Panel . . . 185

7.6 Immunoassays . . . 186

7.7 Censored Data . . . 186

7.8 Statistical Methods . . . 188

7.8.1 Classical Statistics . . . 188

7.8.2 Multivariate Data Analysis . . . 189

8 Synovial Fluid Biology of End-stage Structural Phenotypes 193

8.1 Introduction . . . 193

8.2 Null Hypotheses . . . 194

8.3 Methodology . . . 195

8.3.1 Patient Cohorts . . . 195

8.3.2 Clinical and Radiographic Assessments . . . 196

8.3.3 Biological Assessment . . . 196

8.4 Results . . . 198

8.4.1 Quality Control of SF Analysis . . . 199

8.4.2 Internal Validity of SF Measurements . . . 201

8.4.3 Between-group Comparisons of SF Biology . . . 201

8.4.4 Variability of SF biology . . . 210

8.4.5 Multivariate Analysis of SF Biology . . . 211

8.5 Discussion . . . 215

8.6 Conclusions . . . 219

9 Synovial Fluid Fingerprinting of End-stage Knee Osteoarthritis 220

9.1 Introduction . . . 220

9.2 Null Hypothesis . . . 221

9.3 Methodology . . . 222

9.3.1 Patient Cohorts . . . 222

9.3.2 Clinical and Radiographic Assessments . . . 222

9.3.3 Biological Assessment . . . 223

9.4 Results . . . 224

9.4.1 Quality Control of SF Analysis . . . 225

9.4.2 Between-group Comparisons of SF Biology . . . 226

9.4.3 Categorical Markers . . . 232

9.4.4 Multivariate Analysis of SF Biology . . . 233

9.5 Discussion . . . 242

9.6 Conclusions . . . 261

10 Thesis Discussion **262**

11 Directions for Future Work **271**

A Immunoassay Principles and Performance **276**

A.1 Immunoassay Platforms 276

A.2 Immunoassay Performance Parameters 280

B Materials and Equipment **284**

B.1 Part I Immunoassays 284

B.2 Part II Immunoassays 284

B.3 Materials and Equipment 287

C Limits of Detection and Quantification **288**

D Quality Control **291**

E Reliability of Radiographic Grading **298**

F Part II Supplementary Data **300**

F.1 PLS-DA Regression Diagnostics 300

F.2 Correlation Analysis 305

F.3 Knee Injury and Inflammatory Arthritis Patients 308

Bibliography **310**

List of Figures

1.1 **The normal synovial joint** 4

1.2 **Anatomy of the knee joint** 4

1.3 **Zones of articular cartilage** 11

1.4 **Pathological features of osteoarthritis** 21

1.5 **A new staging paradigm for osteoarthritis** 39

1.6 **Biomarker kinetics** 48

1.7 **Number of osteoarthritis biomarker publications per year** 53

3.1 **Within-subject experimental design for comparison of SF treatments on the Luminex and MSD platforms** 76

3.2 **Polystyrene-bead Luminex: Comparison of bead events & functional samples in neat and HAse treated samples** 80

3.3 **Polystyrene-bead Luminex 10-plex: Comparison of intra-assay precision between neat and HAse treated samples** 81

3.4 **Polystyrene-bead Luminex: Within-subject (pairwise) comparison of dilution-adjusted concentration in functional samples when neat or HAse treated (N=7)** 82

3.5 **Comparison of [A] intra-assay CV, [B] dilution-adjusted IL-6 concentration and [C] analyte recovery in the magnetic bead Luminex IL-6 assay (N=7)** 83

3.6 **MSD 4-plex: Within-subject (repeat measure) comparison of intra-assay precision** 84

3.7 **MSD 4-Plex: Repeat measure comparison of dilution-adjusted concentrations** 85

3.8 **MSD 4-Plex IL-1β assay** 86

4.1 **Within-subject experimental design for comparison and validation of Luminex and MSD platforms** 95

4.2 **Comparison of LOD and Number of samples >LOD for IL-1β, TNF-α, IL-6 and IL-8 between pLuminex, mLuminex and MSD platforms** . 103

4.3 **Comparison of LOQ and Number of samples >LOQ for IL-1β, TNF-α, IL-6 and IL-8 between pLuminex, mLuminex and MSD platforms** . 107

4.4 **Weighted Deming regression of IL-6 measurements by the pLuminex, mLuminex and MSD platform compared to ELISA** 113

4.5 **Bland-Altman plots for agreement of IL-6 measurements by the pLuminex, mLuminex and MSD platforms with ELISA** 114

4.6 **Weighted Deming regression for comparison of IL-6 measurements** 117

4.7 **Weighted Deming regression for comparison of IL-8 measurements** 118

4.8 **Weighted Deming regression for comparison of IL-12 and Eotaxin measurements between Luminex platforms** 119

4.9 **Weighted Deming regression for comparison of MCP-1 and IP-10 measurements between Luminex platforms** 120

4.10 **Bland-Altman plots for agreement of IL-6 measurements between multiplex platforms** 122

4.11 **Bland-Altman plots for agreement of IL-8 measurements between multiplex platforms** 123

4.12 **Bland-Altman plots for agreement of IL-12 and Eotaxin measurements between Luminex platforms** 124

4.13 **Bland-Altman plots for agreement of MCP-1 and IP-10 measurements between Luminex platforms** 125

5.1 **Experimental design for assessing inter-compartment reproducibility of SF Analysis** 143

5.2 **Internal validation of compartment-specific SF analysis using rheumatoid arthritis SF samples (N=8)** 145

5.3 **Cross-compartment coefficient of variation (cCV) for SF measurements in AMOA patients and TCOA patients** 148

5.4 **Compartment-specific concentrations of IL-8, IL-12, TGF-β1 and TGF-β3 in patients with AMOA** 149

5.5 **Compartment-specific TGF-β1, TGF-β3 and IL-12 concentrations in patients with TCOA** 150

5.6 **Proportion of samples measurements above LOD and LOQ by compartment in AMOA and TCOA patients** 152

6.1 **Experimental design for assessing short-term temporal stability of SF analysis** 164

6.2 **Pairwise comparison of Oxford Knee Score (OKS) and Visual Analogue Pain Score (VAS_{pain}) at pre-operative assessment clinic (POAC) and on the morning of surgery** 165

6.3 **Intra-individual SF measurements for IL-6, IL-8, IL-12, IL-15, Eotaxin and MCP-1 at preoperative assessment (POAC) and then at Surgery** 169

6.4 **Intra-individual SF measurements for IP-10, TGF-β1, TGF-β2 and TGF-β3 at preoperative assessment (POAC) and then at Surgery** . . 170

6.5 **Interval coefficient of variation (iCV) for SF measurements** 171

6.6 **Bland-Altman plots for agreement of IL-6, IL-8 and IL-12 measurements taken at pre-operative assessment (POAC) and Surgery** . . . 172

6.7 **Bland-Altman plots for agreement of IL-15, Eotaxin and MCP-1 measurements taken at pre-operative assessment (POAC) and Surgery** . 173

6.8 **Bland-Altman plots for agreement of IP-10 and TGF-β1 measurements taken at pre-operative assessment (POAC) and Surgery** . . . 174

6.9 **Bland-Altman plots for agreement of TGF-β2 and TGF-β3 measurements taken at pre-operative assessment (POAC) and Surgery** . . . 175

6.10 **Proportion of samples measurements above LOD and LOQ by time point** 176

8.1 **Heat map of relative SF analyte concentrations for individual patients** 202

8.2 **Heat map of standardised median SF analyte concentrations by cohort** 203

8.3 **Heat maps for frequency of quantification and detection by cohort for analytes not qualifying for quantitative analysis** 204

8.4 **Comparison of pro-inflammatory cytokines and DcR3 between end-stage knee OA phenotypes** 206

8.5 **Comparison of chemokines and TGF-β isoforms between end-stage knee OA phenotypes** 207

8.6 **Comparison of matrix enzymes between end-stage knee OA phenotypes** 208

8.7 **Comparison of markers of cartilage metabolism between end-stage knee OA phenotypes** 209

8.8 **PCA observation score plots for end-stage knee OA structural phenotypes** 214

9.1 **Comparison of pro-inflammatory cytokines and DcR3 between cohorts** 227

9.2 **Comparison of chemokines and TGF-β isoforms between cohorts** . 229

9.3 **Comparison of matrix enzymes between cohorts** 230

9.4 **Comparison of markers of cartilage metabolism between cohorts** . . 231

9.5 **PLS-DA model using 20 quantitative SF analytes** 235

9.6 **Variable regression coefficient plots for the full PLS-DA model** . . . 237

9.7 **Streamlined PLS-DA model using 8 quantitative SF analytes** 238

9.8 **Variable coefficient plots for streamlined PLS-DA model** 239

9.9 **PLS-DA model using 12 categorical SF analytes** 241

9.10 **Variable coefficient plots for categorical analyte PLS-DA model** . . 242

F.1 **Model diagnostics for full PLS-DA model from 20 quantitative markers** 302

F.2 **Model diagnostics for streamlined quantitative PLS-DA model** . . . 303

F.3 **Model diagnostics for PLS-DA model from 8 categorical markers** . 304

F.4 **Correlation map for end-stage knee OA** 307

F.5 **Correlation map for knee injury** 307

List of Tables

1.1 **Major extracellular matrix constituents in normal articular cartilage** 5

1.2 **Different collagen types in articular cartilage.** 6

1.3 **Causes of secondary osteoarthritis** 19

1.4 **Pathophysiological processes and mediators involved in osteoarthritis** 22

1.5 **Compositional MRI techniques for the assessment of articular cartilage** 34

1.6 **BIPEDS classification of osteoarthritis biomarkers** 45

1.7 **Selected biomarkers currently being evaluated for osteoarthritis** . 52

2.1 **Description of patient cohorts recruited for thesis experiments** . . 65

2.2 **Kellgren & Lawrence radiographic grading system** 67

2.3 **OARSI Atlas radiographic grading system for the knee** 67

3.1 **Recoveries of SF treatment methods (w.r.t neat) for MSD Pro-inflammatory 4-Plex** 85

4.1 **Spiked concentrations (pg/ml) for the mLuminex human cytokine 25-plex assay.** 96

4.2 **Comparison of the number of ideal samples for each assay (bead region) on the mLuminex and pLuminex 25-plex** 101

4.3 **Comparison of the number of functional samples for each assay (bead region) on the mLuminex and pLuminex 25-plex** 102

4.4 **Limits of detection (LOD) for each analyte and platform** 104

4.5 **Proportions of SF samples (N=31) above the limit of detection (LOD) for each analyte and platform given in Table 4.4** 105

4.6 **Limits of quantification (LOQ) for each analyte and platform.** . . . 106

4.7 **Proportions of SF samples (N=31) above the limit of quantification (LOQ) for each analyte and platform given in Table** 4.6 107

4.8 **Spike recoveries for MSD 4-plex assays** 109

4.9 **Spike recoveries for mLuminex 25-plex assays** 109

4.10 **Average replicate CVs as a measure of intra-assay precision** 111

4.11 **Proportion of replicate CVs <20%** 111

4.12 **Concordance correlation coefficient (CCC) and intraclass correlation coefficients (ICC) for absolute agreement and consistency for ELISA and multiplex measurements of IL-6** 112

4.13 **Intraclass correlation coefficients (ICC) for absolute agreement & consistency** 115

4.14 **Concordance correlation coefficients (CCC)** 116

5.1 **Comparison of radiographic assessments between compartments** . 144

5.2 **Intraclass correlation coefficients (ICC) for absolute agreement between compartment-specific analyte measurements** 146

5.3 **Intraclass correlation coefficients (ICC) for absolute agreement between compartment-specific analyte measurements for TCOA patients with and without the potential outlier, patient 12** 146

6.1 **Clinical and radiographic assessments of Experimental group** . . . 166

6.2 **Clinical & radiographic assessments of Reference group** 166

6.3 **Comparison of Experiment (N=10) and Reference patients (N=15) for external validity** 167

6.4 **Intra-class correlation coefficients (ICC) and concordance correlation coefficients (CCC) for absolute agreement between analyte measurements taken at POAC and Surgery** 170

7.1 **Description of patient cohorts recruited for studies in Part II of the thesis** 184

7.2 **Wide-spectrum analysis of synovial fluid: a biological panel of 35 markers** 185

7.3 **Confusion matrix for a 3-group predictive model** 191

8.1 **Description of study cohorts and number of synovial fluid samples in Chapter 8** 195

8.2 **Summary of clinical & radiographic assessments for patients with anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp), tri-compartmental OA (TCOA)** 200

8.3 **Proportion of SF samples above LOQ for analytes not qualified for categorical concentration analysis** 210

8.4 **Proportion of SF samples above LOD for analytes not qualified for categorical concentration analysis** 211

8.5 **Variability of clinical, radiographic and biological endpoints for structural phenotypes of end-stage knee OA** 212

8.6 **Quantitative and Categorical markers used for multivariate analysis** 213

9.1 **Description of study cohorts and number of synovial fluid samples in Chapter 9** 222

9.2 **Clinical and radiographic assessments for study cohorts in Chapter 9** 225

9.3 **Proportion of SF samples above LOQ for analytes not qualified for quantitative concentration analysis** 232

9.4 **Proportion of SF samples above LOD for analytes not qualified for quantitative concentration analysis** 233

9.5 **Quantitative and Categorical markers used for multivariate analysis.** 234

9.6 **Confusion matrix for prediction of end-stage knee OA, knee injury and inflammatory arthritis by quantitative SF analytes** 236

9.7 **Confusion matrix for prediction of end-stage knee OA, knee injury and inflammatory arthritis by streamlined quantitative SF analytes** 239

9.8 **Confusion matrix for prediction of end-stage knee OA, knee injury and inflammatory arthritis by categorical SF analytes** 240

B.1 **Immunoassays used in Part I of the thesis** 285

B.2 **Immunoassays used in Part II of the thesis** 286

C.1 **Analytical limits for Luminex human cytokine 25-plex (Life Technologies, LHC0009M)** 289

C.2 **Analytical limits for other Luminex assays** 289

C.3 **Analytical limits for Meso Scale Discovery assays** 290

C.4 **Analytical limits for ELISAs** 290

D.1 **Isolated anteromedial knee OA (AMOA). N=15** 292

D.2 **Anteromedial knee OA with patello-femoral OA (AMOAp). N=15** . 293

D.3 **Tri-compartmental knee OA. N=30** 294

D.4 **Knee injury. N=20** 295

D.5 **Inflammatory arthritis. N=18** 296

D.6 **Lateral compartment knee OA (Validation test group). N=10** . . . 297

E.1 **Intra-observer (test - retest) reliability for radiographic assessments** 299

E.2 **Inter-observer reliability for radiographic assessments** 299

F.1 **Knee Injury Cohort (N=20)** 308

F.2 **Inflammatory arthritis cohort (N=18)** 309

List of Abbreviations

AMOA Isolated anteromedial knee osteoarthritis

AMOAp Anteromedial knee osteoarthritis with patellofemoral osteoarthritis

CV Coefficient of variation

DMOAD Disease-modifying osteoarthritis drug

ELISA Enzyme-linked immunosorbent assay

esOA End-stage knee osteoarthritis

FWE Family-Wise Error

LCOA Lateral compartment knee osteoarthritis

LOD Limit of detection

LOQ Lower limit of quantification

mLuminex Magnetic bead Luminex assay

MSD MesoScale Discovery

OA Osteoarthritis

OAI Osteoarthritis Initiative

OARSI Osteoarthritis Research Society International

OKS Oxford Knee Score

PCA Principal Component Analysis

PLS-DA Partial Least Squares Discriminant Analysis

pLuminex Polystyrene bead Luminex assay

SF Synovial fluid

TCOA Tri-compartmental knee osteoarthritis

ULOQ Upper limit of quantification

Chapter 1

General Introduction

Osteoarthritis (OA) currently cannot be cured. Is it reasonable to expect to cure a condition that is neither fully understood nor can be reliably measured?

OA is the most common musculoskeletal disorder. It is a leading cause of pain, disability and psychological distress to individuals and places a significant socio-economic burden on healthcare systems and government resources. The knee is the most commonly affected lower limb joint and there are over 250 million people worldwide with knee OA [1]. There are numerous guidelines for the treatment of OA and most recommend a holistic multimodal approach. Joint replacement has been established as a successful intervention for advanced disease refractory to non-operative management, but it is not without its risks, complications and drawbacks.

None of the current treatment modalities for OA address the underlying pathology and most are largely palliative. Despite intense interest from academic institutions and the pharmaceutical industry, there are currently no safe and effective interventions that can preserve joint structure or modify disease progression. The two major obstacles to the development of disease-modifying OA drugs (DMOAD) have been OA's complex pathophysiology and the lack of adequate methods to diagnose and monitor disease activity.

This chapter elaborates on the background and motivations for the thesis. It also acts as a primer for the basic concepts and terms used in the work.

1.1 The Synovial Joint

The synovial (or diarthodial) joint is the organ of movement [Figure 1.1]. It has evolved to allow smooth, low-friction, and pain-free motion between articulating bone surfaces, while providing adequate stability and distribution of load. Despite significant advances in joint arthroplasty, there is no prosthesis that comes even close to matching the function and durability of the synovial joint.

The metaphyses of the articulating bone ends are often expanded and the cancellous bone is capped by a thin layer of condensed cortical bone called the subchondral bone plate. Articular cartilage then covers the subchondral plates of the opposing bone ends. A fibrous capsule encloses the joint and is firmly attached to the bone near the articular surfaces. The capsule is internally lined by the synovial membrane, which is then reflected at the capsular insertions to the margins of the articular cartilage. The synovial membrane defines the intra-articular space (joint cavity) and produces the synovial fluid that bathes the articular surfaces. The joint capsule is reinforced by ligaments, between the juxtaposed bones, that can be extra-articular, intra-articular or both. Some joints contain complete or incomplete fibrocartilaginous discoid structures known as menisci that are fixed to the joint margins via attachments to the joint capsule, ligaments or bone. Joint movement is generated by adjacent muscles via tendons that cross the joint before inserting into bone. The precise morphology of any particular joint is determined by its location and the often competing demands of motion and stability.

1.1.1 The Knee

The knee is the largest joint in the human body [Figure 1.2]. It functions as a hinge allowing flexion and extension, and a small amount of rotation. This allows efficient placement of the foot during gait.

The knee is a combined articulation of the tibia and femur (tibio-femoral joint), and the patella with the femur (patello-femoral joint). The tibial plateau consists of medial and lateral articular surfaces separated by an intercondylar eminence with spinous processes.

The distal femur has 2 convex cam-shaped (medial and lateral) condyles separated by an intercondylar notch. Anteriorly they join to form the trochlea groove that articulates with the patella. The knee is considered to have 3 articular compartments: the medial tibio-femoral, lateral tibio-femoral and patello-femoral compartments. The medial and lateral compartments each contain a crescent-shaped meniscus that is triangular in cross-section. They can be considered as soft tissue extensions of the tibia that improve articular congruity, load transmission and stability. The knee has 4 major ligaments: extra-articular medial and lateral collateral ligaments that provide coronal plane stability against valgus and varus forces; intra-articular anterior and posterior cruciate ligaments that resist translation between the tibia and femur in the sagittal plane. The cruciate ligaments also provide rotational stability.

1.1.2 Articular Cartilage

Arguably the greatest contribution to the exceptional function of the synovial joint is by the articular cartilage that forms its bearing surfaces. This hyaline ("glass-like") cartilage is a highly specialised and resilient load-bearing connective tissue [2]. The thickness of articular cartilage varies between and within human synovial joints, ranging between 2 - 5mm. The cartilage lining the patella is the thickest in the body. Articular cartilage provides a low friction, highly lubricated surface with minimal wear characteristics to allow repetitive gliding motion over several decades.

Articular cartilage consists primarily of a specialised gel-like extracellular matrix (>95% of wet weight) that houses a sparse population of cells called chondrocytes (<5% of wet weight) that manufacture and maintain the matrix. It has no vascular or lymphatic supply, no innervation and is almost non-immunogenic.

1.1.2.1 Extracellular Matrix

The extracellular matrix (ECM) is responsible for the biomechanical and therefore functional properties of articular cartilage. It is a biphasic material comprising tissue fluid and a structural macromolecular framework. The macromolecular framework,

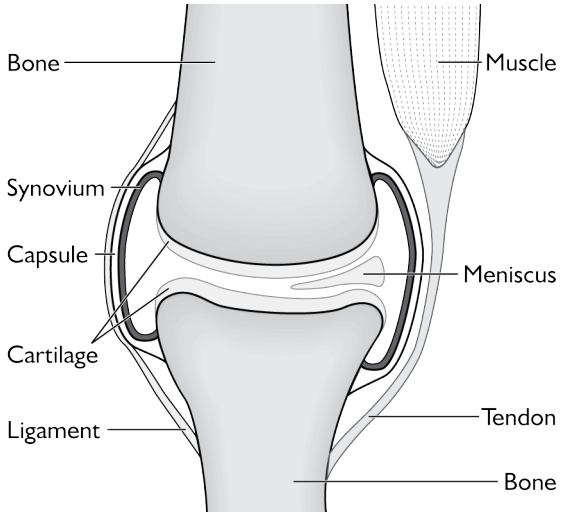


Figure 1.1 – The normal synovial joint

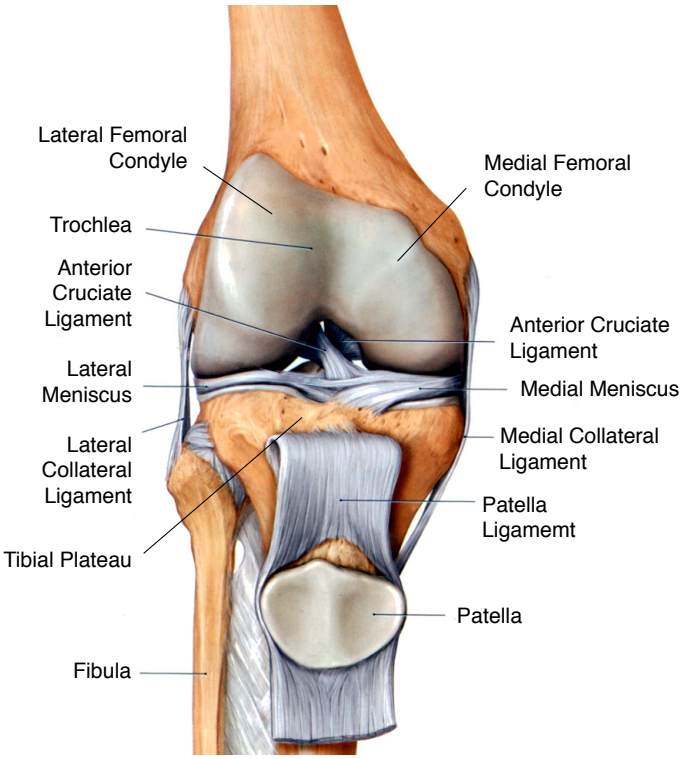


Figure 1.2 – Anatomy of the knee joint. Patella reflected distally.

Component	Wet weight (%)	Dry weight (%)
Water	65 - 80	-
Collagen	10 - 20	60 - 80
Proteoglycan	5 - 10	25 - 35

Table 1.1 – Major extracellular matrix constituents in normal articular cartilage. The predominant collagen is type II collagen and the major proteoglycan is aggrecan.

or "solid matrix", consists mainly of fibrillar type II collagen and the proteoglycan, aggrecan [Table 1.1]. Additional collagens, proteoglycans and non-collagenous proteins and glycoproteins are also present in small amounts. While the solid matrix provides structure and stability, it is the interaction between the tissue fluid and solid matrix that is responsible for the major mechanical properties of articular cartilage.

The ECM also serves as a mechanical signal transducer that can modulate chondrocyte activity, while simultaneously protecting them from excessive mechanical loads. Paracrine communication between chondrocytes by soluble mediators such as cytokines and growth factors occurs through the matrix, which also acts as a reservoir for latent matrix pro-enzymes that can be recruited when necessary.

Collagens: Collagens are the major structural component of the extracellular matrix and contribute 60 - 80% of the dry weight of articular cartilage. The predominant collagen is type II (90-95%) with smaller amounts of other types also contributing to the mature matrix i.e. types VI, IX, X and XI [Table 1.2]. The collagen content is greatest at the cartilage surface. Type II collagen forms a fibrillar meshwork that gives cartilage its form, and tensile and shear properties. The orientation of collagen fibrils within the ECM resists tensile and shear forces on the joint surface and guides them vertically into the subchondral bone. The meshwork also entraps proteoglycans within the matrix and counteracts their swelling pressure.

All collagens exhibit a characteristic triple-helical structure for most of the molecule's length that may be interrupted by one or more non-helical domains. The right-handed triple helix is formed by three polypeptide α chains. Type II collagen molecules (or monomers) of uninterrupted triple-helix assemble head-to-tail and side-by-side in a characteristic quarter-staggered array to form a collagen fibril. Type II collagen fibrils

Collagen Type	Proportion	Function
II	>90%	Major structural fibrillar collagen
III	<10%	Matrix repair and remodelling
IX	1%	Formation and stabilisation of fibrillar meshwork
X	<1%	Structural support and matrix mineralisation
XI	3%	Template for fibril formation
VI	<1%	Chondrocytes attachment to surrounding matrix
XII/XIV	<1%	Non-covalent fibril binding

Table 1.2 – Different collagen types in articular cartilage. The proportion of total collagen content and function for each collagen type are given. Adapted from Eyre et al. [3]

in articular cartilage are found cross-linked to types IX and XI collagen as a functional heteropolymer to facilitate the formation and stabilisation of the fibrillar meshwork.

Proteoglycans: Proteoglycans are large hydrophilic molecules that are principally responsible for the high water content of articular cartilage and confer its compressive strength. Aggrecan is a large aggregating proteoglycan that accounts for ~90% of proteoglycan content. The proteoglycan content of articular cartilage is least at the articular surface, becoming greater and more uniformly distributed with increasing depth.

Proteoglycans consist of one or more glycosaminoglycan (GAG) chains covalently bound to a linear core protein. GAGs are long unbranched polysaccharides consisting of repeating disaccharide units that contain an amino sugar. Each disaccharide unit has at least one negatively charged carboxyl (COO^-) and/or sulphate (SO_3^-) group giving the GAG and subsequent proteoglycan a very high negative charge density. Chondroitin sulphate (4- and 6-isomers) is by far the most abundant GAG found in articular cartilage proteoglycans followed by keratan sulphate and dermatan sulphate. Hyaluronic acid is a non-sulphated GAG and is present as very large unbranched chains, but since it is not covalently bound to a protein core it does not form proteoglycans.

Aggrecan consists of large numbers of chondroitin sulphate and keratan sulphate GAG chains bound to a long protein core. The N-terminal (G1 domain) of the core protein binds non-covalently to hyaluronic acid via a link protein. Vast numbers of aggrecan molecules bind to a single long hyaluronic acid chain to form large proteoglycan

aggregates. Aggregation stabilises the aggrecan molecules and effectively immobilises the proteoglycans within the collagen meshwork, preventing displacement during cartilage deformation. Aggregate size decreases with age and cartilage degeneration.

Small non-aggregating proteoglycans include biglycan, decorin, fibromodulin, syndecan and lumican. These may interact with collagens and aggrecan, influence chondrocyte function and modulate cytokine activities. Lubricin (also known as superficial zone protein or proteoglycan-4) plays an important role in joint lubrication. Although they represent <5% of the proteoglycan mass, the number of molecules (molar amount) is equivalent or greater than aggrecan. Therefore, they may play an important, but yet undefined, role in articular cartilage biology and biomechanics. Cartilage also contains large non-aggregating proteoglycans (< 10%) that may represent degraded aggrecans.

Porous-permeable solid matrix: Proteoglycan aggregates are interspersed and confined within a fibrillar meshwork of collagen type II forming a composite, porous-permeable solid matrix. The negatively charged proteoglycans are trapped and packed to within $1/5$ of their free solution volume within the inter-fibrillar space. Intermolecular friction and steric exclusion limits the expansion of the large proteoglycans through the matrix. This unique collagen-proteoglycan matrix plays a fundamental role in resistance to compressive stresses. The lack of covalent bonds between proteoglycan and collagen allows the collagen fibres to glide through the proteoglycan gel and resist tensile stress.

Non-collagenous proteins and glycoproteins: These represent a diverse group of important molecules with poorly understood functions. They often comprise a protein with a few attached monosaccharides or oligosaccharides. Many of this group may act as “tissue glue”, helping to maintain and organise the macromolecular structure. Other potential roles include chondrocyte-matrix interactions, and modulation of chondrocyte activity. For example, cartilage oligomeric protein (COMP) pentamers binds to both matrix collagen and aggrecan, and chondrocytes. It is thought to facilitate collagen fibril formation, matrix organisation and cell-matrix interactions [4, 5].

1.1.2.2 Tissue Fluid

Water is the most abundant constituent of articular cartilage. It accounts for ~80% of the wet weight at the surface, decreasing to ~65% in the deep zone. Most of the water is contained in the molecular pore spaces created by the collagen-proteoglycan solid matrix. Some water is within the collagen intra-fibrillar space (~30%) and a very small fraction is in the intracellular space. Free inorganic salts, such as sodium, potassium, calcium and chloride, are dissolved within the tissue water.

The high water content of cartilage is largely due to the hydrophilic nature of the proteoglycans. The close proximity of fixed negative charges along the GAG chains attract cations (i.e. Na^+, K^+ and Ca^{2+}), but repel one other and other anions. This unequal distribution of free ions creates a Donnan osmotic pressure that draws water into the matrix. The high negative charge density also results in a strong electro-repulsive force that contributes to the compressive strength of cartilage. The entropic tendency of proteoglycans to gain volume in solution contributes to the total swelling pressure. However, the constraining forces within the strong collagen network surrounding the trapped proteoglycans limit the degree of proteoglycan and therefore cartilage hydration.

The flow of water through cartilage and across the articular surface facilitates nutrient transport and is a mechanism for joint lubrication. Cartilage, however, has a very low permeability. The movement of water through the ECM is limited by high frictional resistance to flow through the molecular pores. Consequently, high pressure gradients or compressive forces are required. These properties allow cartilage to resist compressive stress and support high joint loads.

1.1.2.3 Chondrocytes

Chondrocytes are the only resident cells found in articular cartilage. These highly specialised cells are derived from mesenchymal stem cells and have a very limited proliferative capacity after skeletal maturity. They are truly remarkable in that they alone manufacture, organise, maintain and remodel the complex extracellular matrix

according to the biomechanical needs of the joint while living in an avascular low oxygen environment.

The typical mature human chondrocyte is spherical with a diameter of 10 - 15μm, although cells from different cartilage zones differ in size, shape and probably metabolic activity (chondrocyte phenotypes). However, all chondrocytes possess the capacity and organelles for matrix synthesis and catabolism. They lie within matrix cavities called lacunae and surround themselves with matrix that isolates them from one another. Chondrocytes have no direct cell-to-cell connections, but are intricately connected to their surrounding matrix via fine cytoplasmic processes. These cell-matrix interactions are facilitated by integrins, minor collagens, small non-aggregating proteoglycans and non-collagenous glycoproteins. They also have numerous cell surface receptors for matrix products, cytokines, chemokines and growth factors. Chondrocytes represent the reactive element of articular cartilage. They are able to respond to numerous chemical and physical stimuli including soluble mediators (cytokines and growth factors), matrix composition, static and dynamic mechanical loads, hydrostatic pressure changes and electrical fields.

Adult articular cartilage is avascular and chondrocytes rely on diffusion from the synovial fluid for oxygen, nutrients and metabolite exchange. The cartilage matrix has a small pore size and restricts the diffusion of molecules with respect to molecular weight, size, structure and charge. The diffusion of nutrients and oxygen through the matrix is aided by the movement of fluid during intermittent joint loading and motion. Despite these mechanisms, chondrocytes operate at low oxygen tensions of ~10% at the surface decreasing to 1% in the deeper zones. They are well adapted to these hypoxic conditions and HIF (hypoxia-inducible factor) transcription factors play a key role in cell survival, energy production and matrix synthesis [6]. Because of the low oxygen supply, chondrocytes depend primarily on anaerobic energy production through glycolysis. Individual chondrocytes have a high metabolic activity and their glycolytic rate per cell is equivalent to cells in vascularised tissues. However, the total metabolic rate of cartilage is low due low cell density.

Cartilage Homeostasis: Chondrocytes maintain a stable matrix composition and direct a low rate of turnover and remodelling to meet normal functional demands. Existing matrix is broken down through the activation of matrix enzymes and is replaced through synthesis of new components. Chondrocytes are able to integrate multiple sources of biological and mechanical information from their microenvironment and modulate anabolic and catabolic activity accordingly. In normal cartilage, homeostasis mainly involves the control of proteoglycans, minor collagens and non-collagenous proteins. The half-life of aggrecan is variable between months to years [2]. The type II collagen mesh is extremely stable with virtually zero turnover and normally has a half-life exceeding 100 years [3].

1.1.2.4 Articular Cartilage Structure

Articular cartilage has a unique and highly ordered structure with two levels of organisation: zones and regions. There are 4 zones or layers based on depth from the articular surface and 3 regions or compartments of the matrix defined by the distance from the chondrocyte.

Zones: The surface-to-bone zones vary in collagen fibre thickness and orientation, proteoglycan and water content, and chondrocytes shape, orientation and function [Figure 1.3]. The uppermost superficial layer is made almost entirely of collagen, with both collagen and water concentration decreasing with depth. Conversely, the proteoglycan concentration increases with depth. The diameter of collagen fibres increases from ~20nm in the superficial zone to ~120nm in the deep zone.

1. **Superficial (Tangential) Zone *(10-20%)***: This is the thinnest layer and forms a gliding surface adjacent to the joint cavity. The lamina splendens is a clear acellular sheet of thin collagen fibrils arranged parallel to the joint surface. Beneath this is a cellular layer (1 to 3 cells thick) that has flattened elongated chondrocytes with their long axis parallel to joint surface. The tangential orientation of fibres confers tensile stiffness, resistance to shear stress and allows gliding. The superficial zone matrix has the highest water and collagen content,

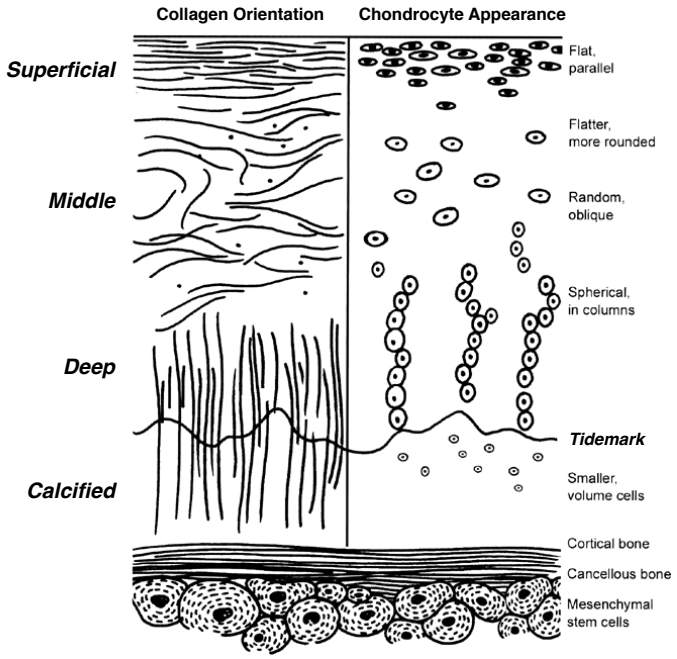


Figure 1.3 – Zones of articular cartilage

but lowest proteoglycan concentration relative to the other zones. Chondrocytes secrete lubricin into the joint cavity that reduces friction and wear. This layer allows diffusion of small molecules such as glucose, but functions as a physical barrier to large molecules moving between the synovial fluid and cartilage. It may also act as an immunological barrier with the chondrocytes displaying macrophage-like feature.

2. **Middle (Transitional) Zone *(40-60%)***: This is the thickest layer of the articular cartilage. Collagen fibres have a larger diameter with a less orderly oblique orientation. The proteoglycan content is greater than the superficial zone, but the collagen and water content is less. This zone allows transition of the shear forces of the superficial zone to compressive forces in the deeper layers. The chondrocytes retain their characteristic spherical shape and have a greater synthetic capacity than the superficial zone.

3. **Deep (Radial) Zone *(30%)***: The deep zone contains the largest diameter collagen fibres, which are orientated perpendicular to the joint surface and subchondral bone. This configuration distributes load and resists compression.

Spherical chondrocytes are arranged in columns perpendicular to the surface. The deep zone has the highest proteoglycan content and the lowest water content.

4. **Calcified zone**: A thin zone of calcified cartilage separates the hyaline cartilage from the subchondral bone. Collagen fibres are vertically orientated and directly coupled to bone collagen fibres. Hypertrophic chondrocytes produce type X collagen and alkaline phosphatase that promote matrix mineralisation with calcium hydroxyapatite crystals. This layer may act as a diffusion barrier between the cartilage and blood vessels within the subchondral bone. The tidemark is an undulating blue (basophilic) upper boundary between the un-calcified deep zone and calcified zone that is visible on histological (haematoxylin and eosin) staining. It represents a calcification front that migrates towards the surface with age and remodelling, leaving behind previous tidemarks. The cement line, formed during growth plate ossification at skeletal maturity, forms the lower border.

Matrix Regions: In addition to the layered architecture, the matrix has a distinct composition and properties depending on the proximity to the chondrocytes.

- **Peri-cellular region:** This forms a thin region ($\sim 2\mu m$) immediately surrounding the chondrocyte that is adherent to the cell membrane. It contains virtually no fibrillar collagen, comprising mainly proteoglycans, non-collagenous molecules (i.e. laminin) and type VI collagen. The relative proteoglycan content is approximately twice that in the inter-territorial region. This region displays both metabolic and mechanical activity. A chondron is defined as the peri-cellular matrix with enclosed chondrocyte(s). In the middle and deep zones, chondrocytes tend to live in chondrons of 2 or more cells that have no inter-cellular junctions but share the same peri-cellular matrix.
- **Territorial region:** The territorial matrix surrounds the peri-cellular matrix of chondrocyte(s). This has a fine fibrillar collagen network that is adherent to the peri-cellular matrix and criss-crosses to form a basket around the cell. In the deep zone, an envelope of territorial matrix surrounds each column of chondrocytes. Cytoplasmic extensions from chondrocytes project into and through the peri-cellular matrix into the territorial matrix. The peri-cellular and territorial regions

couple the chondrocyte membrane to matrix macromolecules and protect the cell from damage during loading and deformation.

- **Inter-territorial region:** This is the largest matrix region by volume and lies between the territorial matrices around chondrons or individual chondrocytes. The large collagen fibrils are not directed around the chondrocytes. Their orientation and size reflect the zone specific structure described previously. The inter-territorial matrix contains the majority of the collagen type II, proteoglycan and water content, and is therefore largely responsible for the biomechanical properties of articular cartilage.

1.1.3 Subchondral Bone

Immediately adjacent to the calcified zone of articular cartilage is a thin endplate of poorly vascularised cortical bone which is supported by cancellous bone. The osteons of the subchondral bone plate are parallel to the joint whereas the trabeculae of the cancellous bone lie perpendicular to the major deforming stresses.

Subchondral bone is a major shock absorber. Although stiffer than cartilage, it is ten times more deformable than cortical bone and provides a pliable bed that protects the overlying articular cartilage from high joint loads. The deep surface of the calcified cartilage keys into the endplate across an irregular undulating surface that serves to resist shearing between the two tissues.

The osteo-chondral junction is considered impermeable and there is no diffusion or molecular communication between bone and cartilage after skeletal maturity. However, this view is being challenged. In particular, there is evidence of fine crossing vascular channels and projections of uncalcified deep zone cartilage passing through the calcified zone into close proximity to subchondral bone [7, 8]. There is an increasing recognition of the functional synergy between cartilage and bone in health and disease.

1.1.4 Synovium

The synovium is a thin sheet of vascularised connective tissue that encloses the joint cavity [9]. Through the production of synovial fluid, it is the primary source of oxygen and nutrition for articular chondrocytes, clears their waste products and provides lubrication for low friction articulation. It also serves as a non-adherent deformable packaging that allows joint movement.

The synovium consists of an inner intimal layer (or synovial membrane) and subintima that lines the joint capsule. The intima is between 20 - 40 μm and comprises 1 to 2 layers of cells within an avascular fine fibrillar matrix. The majority of cells are resident fibroblasts (type-B synoviocytes) with long cytoplasmic processes that form an interconnected meshwork. They secrete hyaluronic acid and control the composition of synovial fluid. There is also a smaller population of bone marrow derived macrophages (type-A synoviocytes) that assist in removal of joint debris and antigen processing. The subintima can be up to 5mm thick. It contains a sparse population of fibroblasts and immune cells in a loose areolar connective tissue with a rich vascular, lymphatic and nervous network. Synovial innervation allows proprioception and pain transmission. The normal synovium has little inflammatory cytokine production and inhibitory molecules predominate [10].

There is no basement membrane between the intimal and subintimal layers. However, the high concentration of hyaluronic acid and meshwork of cellular processes within the intima impede free exchange of large molecules between plasma and synovial fluid. This molecular sieving means that the synovium is functionally a size-selective semipermeable membrane (blood-joint barrier).

1.1.5 Synovial Fluid

Synovial fluid (SF) is a clear straw coloured viscous fluid that bathes the articular surfaces. It acts as a transport medium for cartilage nourishment and waste removal, and lubricates joint motion. A normal joint contains very little SF, which forms a coating over the articular surface of just 10 - 20μm. There is <4ml of SF within the knee, which is the largest synovial joint.

SF is produced by the synovium as an ultra-filtrate of plasma plus hyaluronic acid and lubricin, which are secreted by synovial fibroblasts (type-B synoviocytes). SF is relatively acellular. The composition of small molecules i.e glucose is similar to that of plasma, but larger molecules are excluded by the blood-joint barrier. The protein content of SF is approximately $1/3$ of plasma, with albumin being the most abundant plasma protein.

Hyaluronic acid is the major macromolecular component of SF with a concentration of 2 - 4mg/ml. It is a non-sulphated GAG that is present as very large molecules in polymeric complexes and is responsible for the fluid's viscoelastic non-Newtonian properties. Lubricin, as the name suggests, plays an essential role in joint lubrication although hyaluronic acid may also contribute. SF viscosity decreases over a period of continued low to moderate stress (thixotropic). However, viscosity can increase under rapidly applied high stress (dilatant). An excellent review of the rheological properties of SF is provided by Fam et al. [11].

It is important to note that although SF is produced by the synovium, its overall composition is also acquired from articular cartilage and subchondral bone. This includes products of normal turnover and remodelling of these 3 major joints tissues, matrix enzymes (active and latent) and soluble mediators. Mediators such as cytokines (inflammatory and regulatory), chemokines, and growth factors are involved in cellular communication between articular cartilage, synovium and subchondral bone [12]. They are present in highly regulated low concentrations and are involved in overall joint homeostasis. Therefore, changes in the biological milieu of the joint are reflected in the SF.

1.1.6 Meniscus and Other Structures

Menisci are fibrocartilage structures that are interposed between the articulating surfaces of some synovial joints. They contain small numbers of fibrochondrocytes that manufacture a matrix comprising mainly water and type I collagen fibres. The best characterised menisci are within the knee. By deepening the articular surfaces they allow better joint congruity and therefore stability, and lower contact stresses. Their

ultrastructural organisation and viscoelastic nature also confer load bearing and shock absorbing functions. Menisci are relatively avascular, but often have an innervation that contributes to joint proprioception. A comprehensive review of meniscal structure and function can be found in Rodeo et al. [13].

The capsule, ligaments and tendons around a synovial joint maintain joint integrity, guide motion and provide stability. They are composed primarily of longitudinal arrangements of type I collagen fibres and share a similar architecture. Ligaments are connections between bones, whereas tendons connect muscles to bone. Muscles generate movements via forces transmitted along their tendons. These forces also contribute to the functional stability of the joint. Furthermore, muscles have a huge capacity to absorb energy and therefore act as shock absorbers to limit excessive joint forces. All of these structures possess a rich innervation with sensory and mechanoreceptors, which are essential for protective mechanisms, joint proprioception and dynamic joint stability.

1.2 Osteoarthritis

1.2.1 Epidemiology

There have been countless epidemiological studies reporting on OA according to the involved joint in various populations across the world [14]. What is clear is that it is common, prevalence and incidence increase with age and overall women are affected more than men. OA represents a significant global burden in terms of patient quality of life and healthcare resources [1].

The knee is the most commonly involved joint in the lower limb and the life time risk of developing symptomatic knee OA is nearly 45% [15]. In the UK there are more than 6 million people with painful OA in one or both knees [16].

The exact epidemiological figures for OA depends on how one defines the diagnosis: the patient's clinical symptoms, structural changes that are usually inferred radiographically, or indeed both [17]. OA can have a long asymptomatic period, but radiographs arguably detect late disease. The potential dichotomy between structure and symptoms is a common and highly important theme that permeates through all aspects of OA.

1.2.2 Definition

Our perceptions and understanding of this complex condition have evolved considerably. However, its definition, classification and diagnostic criteria continue to stir considerable debate. A consensus definition for OA was offered by the World Health Organisation and the American Academy of Orthopaedic Surgeons in 1995:

> "OA diseases are a result of both mechanical and biologic events that destabilize the normal coupling of degradation and synthesis of articular cartilage chondrocytes and extracellular matrix, and subchondral bone. Although they may be initiated by multiple factors, including genetic, developmental, metabolic and traumatic, OA diseases involve all of the tissues of the diarthrodial joint. Ultimately, OA diseases are manifested by

> morphologic, biochemical, molecular, and biomechanical changes of both cells and matrix which leads to a softening, fibrillation, ulceration, loss of articular cartilage, sclerosis and eburnation of subchondral bone, osteophytes, and subchondral cysts. When clinically evident, OA diseases are characterized by joint pain, tenderness, limitation of movement, crepitus, occasional effusion, and variable degrees of local inflammation without systemic effects." [18]

1.2.3 Aetiology

OA can occur in normal synovial joints experiencing abnormal biomechanics or in abnormal synovial joint experiencing normal (or abnormal) biomechanics. Traditionally, OA has been aetiologically classified into *primary (idiopathic) OA* that has no obvious identifiable cause and is most commonly encountered in clinical practice, and *secondary OA* that develops as result of antecedent joint pathology or systemic disorder [Table 1.3]. This distinction is becoming increasingly blurred and largely historical, particularly with the increasing number of risk factors being identified that can predispose to "primary" OA. There is therefore a compelling argument that all OA is secondary, but the cause may not yet be apparent [19, 20].

There are numerous potential risk factors for OA that can be broadly divided into genetic, endocrine/metabolic, mechanical and environmental factors. These can be intrinsic or extrinsic to the patient, systemic or local to the joint, and modifiable or invariant. Barring age and female gender, obesity is one of the strongest and best established risk factor for knee OA [21, 22]. Other important risk factors for knee OA include previous joint injury [23, 24, 25] , physical activity [22, 26], malalignment [27, 28] and genetics [29, 30]. The interactions and the mechanisms by which most risk factors are involved in the initiation and progression of OA remains poorly understood. Furthermore, their relative contributions may influence the distribution and nature of joint involvement. A full review of OA risk factors is beyond the scope of this work and readers are referred to the literature [31, 32].

Trauma	**Infection**
Acute joint trauma	Septic arthritis
Intra-articular fracture	
Chronic occupational/recreational overuse	
Meniscectomy	
Instability	**Inflammation arthritis**
Joint dislocation	Rheumatoid arthritis
Ligamentous injuries	Ankylosing spondylitis
Joint hypermobility	Psoriatic arthritis
Ehlers-Danlos Syndrome	Reactive arthritis
Bone & Joint disorders	**Metabolic & Endocrine disorders**
Slipped upper femoral epiphysis	Alkaptonuria (ochronosis)
Legg-Calvé-Perthes disease	Haemochromatosis
Blount's disease	Wilsons's disease
Leg length inequality	Gaucher's disease
Osteonecrosis (avascular necrosis)	Acromegaly
Paget's disease	Hyperparathyroidism
Osteochondritis	Hypothyroidism
Osteopetrosis	
Dysplastic & Hereditary conditions	**Neuropathic Disorders (Charcot joint)**
Developmental dysplasia of the hip	Diabetes mellitus
Multiple epiphyseal dysplasia	Tabes dorsalis
Chondrodysplasias	Leprosy
Stickler's syndrome	Syringomyelia
	Meningomyelocele
Crystal deposition disorders	**Bleeding Disorders**
Gout	Haemophilia
Calcium pyrophosphate deposition	Von Willebrandt's disease
Basic calcium phosphate	(recurrent haemarthrosis)
Haemoglobinopathies	**Miscellaneous**
Sickle cell disease	Kashin-Beck disease
Thalassemia	

Table 1.3 – Causes of secondary osteoarthritis

1.2.4 A New Paradigm

OA is failure of the entire synovial joint as an organ [33, 34]. Like any other organ failure i.e. renal, cardiac, respiratory or hepatic failure, OA is not a single disease entity. It comprises a heterogenous spectrum of disorders and risk factors, which produce a loss of normal homeostasis that ultimately leads to the functional insufficiency of the synovial joint. OA may begin in any one or more tissues of the synovial joint, but subsequently involves all components, albeit to a variable and often characteristic extent. Furthermore, like other organ failures, there is usually a clinically silent asymptomatic period.

In keeping with the analogy of organ failure, OA is not a disease and is therefore not technically a diagnosis. Instead, its clinical features represents a syndrome. However, this is a semantic issue and the “disease” of OA and its “diagnosis” will be used freely within this thesis.

1.2.5 Pathophysiology

The pathogenesis and pathophysiology of OA is extremely complex [reviewed in 35, 36, 37, 38, 39].

OA is a mechanically driven, but biologically mediated process that is modulated by a complex interplay of genetic, biomechanical, endocrine/metabolic and environmental risk factors. The process involves a dysfunctional repair and remodelling response involving all tissue of the synovial joint in concert, which communicate with each other and can influence each other’s behaviour. There are significant cross-talk and feedback mechanisms particularly between articular cartilage, subchondral bone and synovium. These develop into interconnected vicious cycles that disrupt the normal equilibrium between synthesis and degradation. The loss of joint homeostasis leads inexorably to the biomechanical and biological failure of the entire joint.

Why and how the final common pathway of joint failure is initiated is poorly understood. In particular, the exact nature of the tipping point remains elusive. The characteristic pathological features of OA are shown in Figure 1.4.

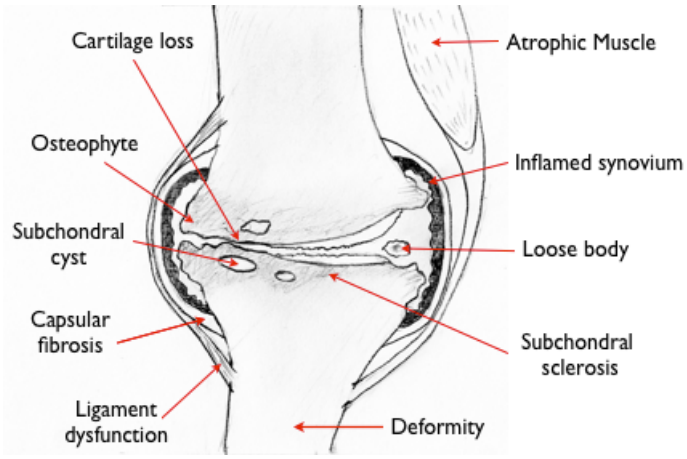


Figure 1.4 – Pathological features of osteoarthritis

1.2.5.1 Matrix Enzymes and Soluble Mediators

The pathophysiology of OA is mediated by a large orchestra of matrix enzymes, growth factors, cytokines, chemokines, eicosanoids and reactive oxygen and nitrogen species; conducted by epigenetic and transcription regulation [Table 1.4] . These mediators and processes can be broadly categorised as either anabolic or catabolic. A brief summary of soluble mediators relevant to this work follows.

Matrix Enzymes: The matrix metalloproteinases (MMPs) and a disintegrin and metalloproteinase with thrombospondin motif (ADAMTS) families of zinc-containing, calcium-dependant proteolytic enzymes are involved in matrix degradation in all connective tissues. [40, 41, 42]. They are produced by chondrocytes and synoviocytes as latent pro-enzymes (pro-MMPs and pro-ADAMTS). Activation of these zymogens is therefore an important regulatory step in their activity. Cartilage proteoglycan is predominantly broken down by aggrecanases (ADAMTS-4 and ADAMTS-5) and MMP-3 (stromolysin-1). Type II collagen is broken down predominantly by MMP-1 (collagenase-1) and MMP-13 (collagenase-3). MMP-13 is the most potent type II collagenase and has received much attention as a pivotal mediator in OA [43]. MMP-3 also activates matrix pro-enzymes and degrades collagen byproducts and type IX collagen. The expression, release and activity of MMPs and aggrecanases are modulated by inflammatory cytokines and reactive oxygen species, and are therefore noted to be greatest in regions of damaged cartilage.

Biologic process	Proposed mediators	Potential therapeutics
Matrix degradation	MMPs 1, 3, 9, and 13, ADAMTS 4 and 5, cathepsin K, serine proteases (HTRA-1) driven by cytokines (ILs 1, 6, 7, 8, 17, and 18, OSM), chemokines (IL-8, GROα, GROγ, RANTES, MCP-1) and others (S100 proteins, TGFα, matrix fragments, leukotrienes, and prostaglandins)	Protease inhibitors, TIMPs, anticytokine therapy, TLR inhibition, MAP kinase inhibition, NF-κB inhibition, lipoxygenase and cyclooxygenase inhibitors
Reduced matrix repair	$\downarrow$ activity of IGF-1, TGFβ, BMP-7 (OP-1), FGF-18	Growth factors (IA or by gene therapy)
Cell death	$\downarrow$ HMGB-2, $\downarrow$ autophagy, reactive oxygen and nitrogen species	Caspase inhibitors, antioxidants, iNOS inhibitors
Chondrocyte hypertrophy	RUNX-2, HIF-2α, Wnt/β-catenin, IL-8	PTH, calcitonin
Calcification and crystals	Transglutaminase, inorganic pyrophosphate, TLRs, NLRP3	Phosphocitrates, TLR and NLRP3 inhibition
Subchondral bone sclerosis	Wnt/β-catenin, $\downarrow$ sclerostin, BMPs, IGF-1	Wnt or BMP antagonists, slowing bone remodeling with bisphosphonates or anti-RANKL
Osteophyte formation	TGFβ, BMP-2	Since these may stabilize the joint, should probably not be targeted directly
Focal bone remodeling (bone marrow lesions)	RANKL, VEGF	Bisphosphonates, anti-RANKL
Synovitis	IL-1β, TNFα, IL-17, IL-15, IL-7, CCL19, MCP-1, MIP-1β, S100 proteins/alarmins	Anticytokine therapy, TLR antagonism, complement inhibition

* MMPs = matrix metalloproteinases; ILs = interleukins; OSM = oncostatin M; GROα = growth-related oncogene α; MCP-1 = monocyte chemotactic protein 1; TGFα = transforming growth factor α; TIMPs = tissue inhibitor of metalloproteinases; TLR = Toll-like receptor; IGF-1 = insulin-like growth factor 1; BMP-7 = bone morphogenic protein 7; OP-1 = osteogenic protein 1; FGF-18 = fibroblast growth factor 18; IA = intraarticular; HMGB-2 = high mobility group box protein 2; iNOS = inducible nitric oxide synthetase; RUNX-2 = runt-related transcription factor 2; HIF-2α = hypoxia-inducible factor 2α; PTH = parathyroid hormone; VEGF = vascular endothelial growth factor; TNFα = tumor necrosis factor α; MIP-1β = macrophage inflammatory protein 1β.

Table 1.4 – Pathophysiological processes and mediators involved in osteoarthritis. Reproduced from Loeser et al. [39].

The activity of these proteolytic enzymes is regulated by tissue inhibitors of metalloproteinases (TIMPs), and proteases such as plasmin and cathepsins. There are 4 known species of TIMPs (TIMP-1, -2, -3 and -4) which inhibit the MMPs with varying degrees of affinity, but TIMP-3 is also a potent inhibitor of aggrecanases [44]. Plasmin and cathepins play a role in pro-enzyme activation.

Cytokines: The role of cytokines in OA pathophysiology is well documented and has received great attention [reviewed in 45, 46]. Cytokines are produced by chondrocytes and synoviocytes, and act in an autocrine and paracrine fashion. They are often classified according to their predominant effects as pro-inflammatory (catabolic), inhibitory (anti-catabolic/anabolic) and regulatory.

Inflammatory cytokines are critical mediators in the disturbed metabolism and enhanced catabolism of OA [47]. Interleukin (IL)-1β, complemented by tumour necrosis factor (TNF)-α, seems to be the main pro-catabolic cytokine. IL-6, IL-8 (also a chemokine) and leukaemia inhibitory factor may not directly stimulate cartilage destruction, but probably modulate the effects of catabolic cytokines. Inflammatory stimulation of chondrocytes induces phospholipase A2, cyclo-oxygenase-2 and inducible nitric oxide synthetase. Their products i.e. prostaglandin-E and nitric oxide may modulate chondrocyte matrix synthesis and degradation in a positive or negative manner.

Inhibitory cytokines such as IL-4, IL-10 and IL-1 Receptor Antagonist (IL-1RA) have anti-catabolic functions and are therefore chondroprotective. These cytokines reduce the production and/or downstream activity of inflammatory cytokines, and can attenuate degradative matrix enzymes and enhance TIMPs.

Chemokines: Chemokines are mediators that closely resemble cytokines and have both potent inflammatory/catabolic and cell trafficking functions [48]. Their release by chondrocytes and synoviocytes is strongly stimulated by inflammatory cytokines [49]. They act on specific cell surface receptors and they function in an autocrine/paracrine manner. Chemokines such as monocyte chemoattractant protein 1 (MCP-1), macrophage inflammatory protein (MIP)-1α, MIP-1β and RANTES (regulated on activation,

normal T cell expressed and secreted) contribute to cartilage catabolism. Actions include reduced proteoglycan synthesis; induction of MMPs (i.e. MMP-1, MMP-3 and MMP-13), cathepsins and lysosomal enzymes; and promotion of chondrocyte apoptosis [50, 51, 52, 48].

Growth Factors: Growth factors such as transforming growth factor (TGF)-β isoforms, bone morphogenetic proteins (BMPs), insulin-like growth factor (IGF)-1 and fibroblast growth factors (FGFs) are the major anabolic force in OA [53, 54, 55]. They are released by chondrocytes, osteocytes and synoviocytes in response to catabolic signals such as matrix degradation and inflammatory cytokines, and biomechanical signals.

Growth factor activity is both directly anabolic and anti-catabolic: they stimulate chondrocyte collagen and proteoglycan synthesis; inhibit matrix degradation by down-regulating proteolytic enzymes and upregulating TIMPs; and reduce production and receptors to inflammatory mediators such as IL-1β. They also promote chondrocyte and osteoblast survival and proliferation. The BMPs and TGF-β play a pivotal role in subchondral bone remodelling and in particular osteophyte formation [56]. Some growth factors such as TGF-β and BMPs can have a complex dual role in OA pathophysiology [54, 57].

Reactive Species: Reactive oxygen species (ROS) and reactive nitrogen species (RNS) play an important role in normal cartilage homeostasis and OA [58, 59]. They include superoxide anions ($^{\bullet}O_2^-$), hydroxyl radicals ($^{\bullet}OH$), nitric oxide ($^{\bullet}NO$) and their derivatives. ROS and RNS production is increased by mechanical stress and pro-inflammatory cytokines. Saturation of intrinsic anti-oxidant mechanisms produces a state of oxidative stress that promotes chondrocyte death, matrix degeneration and further increases pro-inflammatory cytokines.

1.2.5.2 Articular Cartilage

OA is a war fought on many fronts between anabolism/synthesis and catabolism/degradation in which the latter eventually emerges as the victor. Although there are important

influential battles in subchondral bone and synovium, arguably the main frontline remains articular cartilage and it has therefore received the most investigation.

Pathological changes occur at the articular surface where mechanical stresses, and in particular shear stresses are greatest [60]. These include discolouration, softening, roughening, fibrillation, pitting, erosion (ulceration), delamination and eventually denudation to subchondral bone. Overall loss of cartilage thickness is also due to advancement of the calcified cartilage zone at the osteochondral junction.

Matrix Changes: Matrix loss is due to a combination of biological process and mechanical wear [61]. The inability to replicate initially the quality, and eventually the quantity of ECM, predisposes to further damage under increasingly lower mechanical stresses. Biochemical changes in the ECM predominantly occur in the inter-territorial matrix and start in the superficial zone around chondrocytes, extending deeper with progression.

- **Loss of proteoglycans and reduced aggregation:** Loss of the large aggregating proteoglycan, aggrecan, is a consistent and early biochemical finding. There is a rapid loss of proteoglycan that progresses with disease severity. Proteoglycan depletion is due to a combination of leaching from the damaged collagen network and enzymatic breakdown by MMP-3 and aggrecanases (ADAMTS-4 and ADAMTS-5). There is a compensatory increase in chondrocyte proteoglycan synthesis, particularly in early OA. This is associated with altered GAG composition and as OA progresses, the proteoglycan chain lengths become shorter. Aggregation is reduced due to proteolytic activity at the hyaluronic acid binding region of the core protein.

- **Reduced quality and integrity of the collagen network:** Early aggrecan depletion exposes the collagen meshwork. The ultrastructure and organisation of the type II collagen mesh becomes significantly disturbed. Cross-linking is reduced and fibre size and orientation become variable and disordered. Collagenases such as MMP-1 and in particular MMP-13 then break down the type II collagen fibres. This may be the pivotal point of no return in OA. Type II collagen synthesis by

chondrocytes is upregulated and collagen content may be maintained in early disease. However, a progressive reduction in both quality and content of fibrillar collagen inevitably ensues. There is increased and altered metabolism of minor collagens such types III, VI and X and the appearance of “foreign” collagens such as type IIA and even type I.

- **Increase in water content**: An increase in water content is one of the earliest detectable changes in OA, albeit by only a few per cent over normal cartilage. OA cartilage always swells (gains water) despite the loss in proteoglycan. This is due to the increased porosity of the weakened type II collagen meshwork. The collagen network is less able to resist the swelling pressure of existing proteoglycans as they imbibe water.

- **Increased mineralisation of the deep cartilage:** The osteochondral junction reverts to a developmental state akin to physeal growth and endochondral ossification. Deep zone chondrocytes become hypertrophic, secreting MMP-13, type X collagen and alkaline phosphatase, and there is associated matrix mineralisation. The zone of calcified cartilage thickens and advances into the deep cartilage zone effectively reducing articular cartilage thickness. This is noted by replication of the histological tidemark.

The biochemical and structural alterations in the matrix of OA cartilage result in inferior mechanical properties compared to normal articular cartilage. In particular, resistance to compression is reduced and there is failure of load bearing function, which predisposes to and drives further degradation. The inferior quality of the type II collagen network results in reduced tensile stiffness and strength of the solid matrix. Disruption of the solid matrix leads to increased water content (porosity) and a significant increase in permeability. Consequently, there is a loss of interstitial fluid pressurisation that normally provides a significant component of the load support in compression. Load support shifts more to the solid matrix predisposing to further damage by fatigue wear. The increase in permeability also allows loss of lubricant leading to greater interfacial wear.

Chondrocyte Response : Chondrocytes respond to physical and biological changes in their microenvironment i.e. mechanical load (static and dynamic), osmotic pressure, matrix composition and soluble mediators (see Section 1.2.5.1). These relatively quiescent cells are spurred into activity in an attempt to repair and remodel their surrounding matrix. They display a multitude of molecular responses involving autocrine/paracrine cellular signalling via soluble mediators, altered transcriptional and epigenetic modifications, feedback loops involving matrix elements and cross-talk with synovium and subchondral bone [reviewed in 62, 35]. Chondrocyte activity in OA can be categorised broadly into:

- **Proliferation or death:** Chondrocytes undergo limited proliferation in areas of matrix damage. This is an attempt to increase cartilage synthetic activity and compensate for cell loss. Proliferation is mediated by growth factors such as TGF-β and IGF-1. Chondrocyte cloning and cluster formation is most evident in the more superficial cartilage zones. Alternatively, chondrocytes may die by various processes including apoptosis or "chondroptosis" [63, 64]. Elimination of the only source of matrix synthesis does not make biological sense. However, oxidative stress by ROS produced in response to mechanical injury may promote this pathway.

- **Matrix synthesis/catabolism:** Chondrocytes increase synthesis of proteoglycan and collagen, and production of degradative matrix enzymes. The upregulation of both anabolic and catabolic activity is an attempt to repair and remodel the matrix. Anabolic activity is favoured by growth factors, whereas catabolic function is driven by inflammatory cytokines, chemokines and ROS/RNS. Chondrocytes in OA may become hypo-responsive to growth factors. For example, increased levels of IGF-1 binding proteins reduce the bio-availability of IGF-1 [65]. Tipping of the balance to catabolism lies at the core of OA.

- **Phenotypic modulation:** Chondrocytes display phenotypic plasticity and have been reported to assume one of several phenotypes. *Mature* cells continue to synthesise aggrecan and type II collagen matrix components, and can be either predominantly anabolic or catabolic. *De-differentiated* chondrocytes lose this ability and instead produce type I, III and V collagen, serving a wound healing

fibroblast-like function. *Hypertrophic* chondrocytes are *chondroprogenitor* cells that represent states of developmental regression. Hypertrophic cells produce type X collagen, MMP-13 and alkaline phosphates for matrix calcification before undergoing apoptosis. Chondroprogenitor cells express fetal type IIA collagen and differentiation markers. The mechanisms and rationale for these phenotypic shifts is under continued investigation. There is evidence that phenotype may depend on the cartilage zone and therefore the biological and biomechanical microenvironment.

1.2.5.3 Subchondral Bone

Subchondral bone remodelling with sclerosis, cyst and osteophyte formation are characteristic pathological features. The articular cartilage and subchondral articular bone form a functional unit and there is emerging evidence for significant molecular cross-talk during OA [7, 66, 67]. Subchondral bone changes begin very early in the disease process and evolve simultaneously with cartilage changes. There is even a suggestion that changes may precede and induce cartilage pathology, although this remains controversial [68, 69]. OA is associated with increased vascularity of subchondral bone and penetration of vessels through the cement line into the calcified cartilage zone [70]. This is associated with sensory nerve fibres that represent a potential source of pain in OA [71].

Subchondral bone remodelling occurs due to a combination of mechanical loading (Wolff's law), micro-fractures and soluble mediators [reviewed in 72]. Increased mineralisation with thickening of the cortical endplate plate and modified trabecular architecture leads to subchondral sclerosis and eburnation. The increase in subchondral bone stiffness reduces the protective shock absorbing capacity, predisposing the overlying cartilage to damage from further loading. Stress shielding of the metaphsyeal bone results in loss of bone density. As remodelling becomes increasingly abnormal and disordered, the subchondral bone undergoes attrition and collapses under compression leading to deformity.

Subchondral bone cysts represent microfractures of the subchondral trabecular bone at various stages of healing and remodelling with gelatinous replacement of the marrow [73]. Furthermore, synovial fluid may also be forced into the degenerating microfractures through cartilage fissures and an increasingly permeable solid matrix. Bone marrow oedema-like lesions (BMLs) are non-cystic, ill-defined areas of high intensity adjacent to the subchondral bone plate on MRI (T2 or proton density-weighted) images [reviewed in 74]. BMLs are observed in both early and established OA, and in particular knee OA. They correlate spatially with areas of increased cartilage degeneration, bone remodelling, subchondral cyst formation and subsequent bone attrition. The pathogenesis and natural history of BMLs is controversial, but there is evidence that they are predictive for pain and progression in knee OA [74].

Molecular mechanisms involved in subchondral remodelling include inflammatory cytokines; TGF-β and BMP growth factors; the osteoprotegerin (OPG)/receptor activator of nuclear factor-kappaB (RANK)/RANK ligand (RANKL) system; and the Wnt signalling pathway [67].

Osteophytes are a cardinal feature of OA. These develop at the joint margin, adjacent to areas of increased joint loading, and may be an attempt at maintaining joint stability [56]. Proliferation of periosteal or synovial mesenchymal cells differentiate into chondrocytes and take part in endochondral ossification to produce an enlarging bony protrusion. Osteophytes grow in the path of least resistance guided by capsular traction. Mechanical factors, subchondral bone vascularisation and local production of TGF-β, BMP-2, BMP-7 and IGF-1 are implicated in osteophyte formation.

1.2.5.4 Synovium

The relevance of the synovium to OA was not previously established, but is now receiving a great deal of attention. There is strong evidence for the role of synovial inflammation in the pathogenesis and progression of OA [reviewed in 75, 76, 77].

Magnetic resonance imaging (MRI) and arthroscopic assessments have revealed that synovial inflammation is an important feature of OA [78]. The degree of synovitis in OA is highly variable in terms of distribution and intensity. Focal areas of synovitis are

initially found in proximity to pathologically damaged cartilage and bone. Histologically, there is lining or villous hyperplasia, increased vascularity and perivascular infiltration by macrophages and lymphocytes. The overall prevalence and severity of synovitis increases with advancing OA, but there is evidence that it is also an important feature of early disease [79, 80, 81, 82]. The presence and degree of synovitis has been shown to be associated with OA progression [83, 82].

Synovitis is associated with patients' symptomatic severity in OA [84, 85, 86]. In retrospect, this may not come as a huge surprise since many OA symptoms i.e. stiffness and swelling (effusion) can be attributed to inflammation. Furthermore the synovium has sensory innervation that is a potential source of OA pain. This tissue may therefore offer some explanation to the apparent discord between radiographic assessments of OA and symptoms (see Section 1.2.9).

The synovium is an extremely reactive tissue. Inflammation is thought to be initiated as a response to products of cartilage matrix catabolism released into the synovial fluid [reviewed in 77]. This is probably intended as a mechanism to clear debris and restore joint homeostasis. Danger signals activate synovial macrophages and the complement pathway resulting in the production of inflammatory cytokines, chemokines and degradative matrix enzymes. These soluble mediators spill over into the synovial fluid and contribute to chondrocyte dysregulation, tipping the balance in favour of matrix catabolism over anabolism. This promotes further synovitis and a vicious cycle is formed.

1.2.6 Clinical Features and Diagnosis

Patients with OA present with joint pain, swelling and stiffness, loss of mobility and function, and a reduced quality of life [87, 88]. There may be mechanical symptoms in the presence of loose bodies, such as large cartilage fragments, or meniscal pathology. Although generally progressive, these symptoms, and in particular pain, can fluctuate with periods of stability and exacerbation. Physical examination often reveals a joint effusion, bony enlargement and deformity, malalignment, tenderness to palpation and a painfully reduced range of joint motion with crepitus. Radiographic assessment is the

gold standard investigation. The classic hallmarks are joint space narrowing, marginal osteophytes, subchondral bone sclerosis and subchondral bone (pseudo)cysts. This pattern of features combine to give the “syndrome of OA”.

There are several formal diagnostic criteria and recommendations [89, 90]. However, most clinicians would probably admit to taking a less stringent approach in their routine practice. Clinical findings alone are usually sufficient to make a diagnosis, but are nonetheless usually combined with radiographic assessment [91, 90]. The presence of risk factors elicited in the patient’s history can further add to diagnostic confidence. There is evidence that reliance on radiographic criteria favours the diagnosis of more severe disease at the risk of missing earlier disease [92]. Laboratory tests are currently rarely used or required except in excluding other joint pathologies i.e. inflammatory or septic arthritis.

1.2.7 Assessment of Symptoms

The most important symptom of OA is pain. It is the reason why patients seek medical help and has the greatest impact on function and quality of life [87]. Arguably, pain has the greatest influence on the patient’s and surgeon’s decision for operative intervention [93, 94, 95]. Pain is classically described as a constant background ache with sharp exacerbations due to activity. However, it is well accepted that OA pain is highly variable and episodic [88, 96, 97, 98]. Furthermore, OA pain is influenced by a multitude of factors including biological, mechanical, psychological and social [99, 100, 101, 102]. The perception and experience of OA pain is an extremely complex phenomenon and readers are directed to some excellent reviews of this topic [103, 104, 105].

There are a large number of tools available to assess patients’ symptoms [106, 107]. Some focus primarily on pain i.e. visual analogue scale [108] and Intermittent and Constant Osteoarthritis Pain Score (ICOAP) [109]; while others also encompass function and disability i.e. Western Ontario and McMaster Universities Arthritis Index (WOMAC) [110]; Knee Injury and Osteoarthritis Outcome Score (KOOS) [111] and Oxford Knee Score (OKS) [112, 113]. However, many of these instruments were designed as patient reported outcome measures (PROMS) to assess health status of patient cohorts

usually before and after an intervention i.e. joint replacement surgery or time interval. Therefore, their appropriateness in the assessment, staging and monitoring of knee OA symptoms, especially in an individual patient, may not be validated. Furthermore, the minimal clinically important difference for many of these scoring systems remains undefined.

1.2.8 Assessment of Structure

Current methods for diagnosing and monitoring OA are on the whole inadequate. There is a real need for tools that can detect early OA, are sensitive to longitudinal change, correlate with clinical symptoms and that can be used to guide treatments.

1.2.8.1 Radiographs

Radiographs remain the gold standard investigation for OA. They are safe, cheap, fast and easily available. The most common images taken of the knee are weight-bearing anterior-posterior and lateral views and a skyline axial view for the patello-femoral articulation.

There are a number of grading systems used to assess radiographic severity. The Kellgren and Lawrence (KL) scale was first published nearly 60 years ago, yet remains by far the most commonly used grading system in both the literature and clinical practice [114]. A KL grade of 2 is often used as a threshold for the diagnosis of established OA. However, a KL grade of 1 has been shown to represent OA based on risk of progression [115]. The original description was not explicit and its interpretation has been variable resulting in potential problems when comparing studies [116, 117]. Furthermore, the KL system is heavily weighted to osteophytosis and therefore bone remodelling, and may be insensitive to change [118]. The Osteoarthritis Research Society International (OARSI) score is an atlas based grading system [119]. This system makes separate assessments of joint space narrowing and osteophytosis in each of the tibio-femoral compartments and comments on subchondral sclerosis and bone attrition. The patello-femoral compartment is not included in the assessment,

despite this being a highly prevalent and symptomatic area of disease [120, 121, 122]. The OARSI scoring system may be more sensitive to longitudinal change than the KL system [123]. Although the level of detail allows versatility regarding feature and compartment specific subscores, it also makes the system rather cumbersome. There are a number of automated computer assessment methods for knee OA, but these are complex, time consuming and not widely available [124, 125, 126].

Radiographs have several important limitations. They are essentially 2-dimensional images of bone, although they can detect any other calcified structures i.e. loose bodies and calcified menisci, and may offer information on joint alignment. However, they are only an indirect measure of cartilage thickness and provide no information on the soft tissues of the synovial joint i.e. synovium. Joint space width or narrowing (JSW or JSN), which is used to infer cartilage loss, is affected by subtle variations in image projection and joint position [127, 128, 129, 130]. Although generally acceptable for crude appraisal in the clinic, standardised reproducible image acquisition becomes increasingly important for more detailed and longitudinal measurements. The radiographic features of OA represent established pathology and are insensitive to early disease. Changes in radiographic parameters over time are slow and variable. Only 6% (±5%) of people demonstrate evidence of progression over one year using the KL grade and the annual change in JSW was 0.13 ± 0.15 mm/year [131]. The clinical significance of such changes is unknown. Most importantly, JSW is significantly influenced by meniscal pathology and position in the knee and may be an unreliable measure altogether [132].

1.2.8.2 Magnetic Resonance Imaging (MRI)

MRI offers multi-planar, high resolution, and direct visualisation of all tissues of the synovial joint. In knee OA, MRI is commonly used to assess cartilage volume and thickness, synovitis and synovial fluid effusions, BMLs, meniscal pathology and osteophytes. Less frequent assessments include muscle, bursae, ligaments, fat pad and bone defects.

MR Imaging Technique	Biochemical Component Assessed
dGEMRIC	Glycosaminoglycans
T1ρ imaging	Collagen network & glycosaminoglycans
T2 mapping	Collagen network & water content
Sodium imaging	Glycosaminoglycans
Diffusion-weighted imaging	Collagen network & glycosaminoglycans

Table 1.5 – Compositional MRI techniques for the assessment of articular cartilage

Synovial joint evaluation by MRI is described as either morphological (semiquantitative and quantitative) or compositional. Semiquantitative morphological scoring systems include the whole organ MRI score (WORMS), Boston Leeds osteoarthritis knee score (BLOKS) and MRI osteoarthritis knee score (MOAKS) [reviewed in 133]. In these scoring systems, the morphologic characteristics of cartilage are assessed in conjunction with additional structures such as menisci, subchondral bone, osteophytes, and synovium. Quantitative techniques at present are focussed on cartilage thickness and volume [reviewed in 134]. Optimum imaging of synovitis uses contrast-enhanced protocols and are largely semiquantitative, but quantitative methods are available [80, 135]. Physiological, also known as compositional, techniques provide information on the biochemical properties of articular cartilage. These require advanced MRI protocols and include delayed gadolinium-enhanced MRI of cartilage (dGEMRIC), T1ρ imaging, T2 mapping, sodium imaging and diffusion-weighted imaging [Table 1.5] [reviewed in 136, 137]. Compositional MRI gives the opportunity to detect and monitor early, pre-morphological changes.

At present, the clinical utility of MRI evaluation in knee OA beyond the research setting is in the early stages of development. MRI features of OA can be present in ~90% of people without radiographic features of knee OA whether they have pain or not [138]. Whether the changes in asymptomatic individuals represents incidental findings without clinical relevance or true pre-symptomatic OA awaits results from longitudinal studies such as the Osteoarthritis Initiative (OAI) [139]. Although MRI can detect knee OA with high specificity and moderate sensitivity compared with reference standards i.e. radiography, arthroscopy and histology, the sensitivity of MRI is below clinical diagnostic standards [140]. Therefore, it may be more useful in ruling out OA. A potential explanation is the lack of a standard MRI definition for knee OA and there

is now an OARSI consensus definition for use in future studies [141]. However, the reliability of MRI assessments and responsiveness to structural joint changes is excellent and superior to plain radiographs [142, 143].

MRI has a number of technical and practical limitations that preclude its use for OA in clinical practice. It is an expensive investigation that is not readily available. Imaging is complex requiring specific protocols for optimum imaging of different tissues. Movement artefact can significantly degrade quality and therefore usefulness. The examination time is lengthy, both in image acquisition and reporting that requires sophisticated software and trained experts. Some patients may not tolerate the lengthy investigation due to anxiety and claustrophobia, while a high BMI may exclude others. The use of contrast for some MRI protocols i.e. synovitis assessment and dGEMRIC brings with it risks of allergy and nephrotoxicity. Although not as critical as for radiographs, patient positioning remains important especially for repeat studies.

MRI is a powerful and rapidly evolving research tool and its use for structural joint assessment is gaining considerable momentum. Furthermore, resolution continues to increase with more powerful magnets and optimised protocols. Consequently, MRI now serves as a potential source of "dry" biomarkers for OA and will most likely gain regulatory approved for use in clinical trials.

1.2.8.3 Other Modalities

Computer Tomography (CT): CT provides multi-planar 3-dimensional imaging of bone morphology and can detect subchondral bone changes [reviewed in 123]. However, it does not provide any direct information on cartilage or other soft tissues, unless combined with invasive contrast arthrography that is insensitive to early changes. The high radiation exposure, relative expense and lack of real merit beyond radiographs and MRI limits the clinical and research utility of CT.

Ultrasound (US): US can offer dynamic multi-planar imaging of joint soft tissues, and is particularly good for assessing synovitis and effusion [reviewed in 123]. Combination with Doppler can provide information on synovial vascularity and therefore

inflammation. It is routinely available, relatively cheap, imparts no radiation and can be used to guide therapeutic injections. US shows great potential for the evaluation of small joint OA i.e. the hand and foot. However, it lacks adequate penetration to assess large joints and its role in this context is unclear.

Arthroscopy: Although not an imaging modality, arthroscopy offers direct visualisation of structural changes in cartilage, synovium and internal structures such as menisci. Evidence of osteophytosis can also be detected, but subchondral bone changes are otherwise not accessible. Arthroscopy only allows macroscopic inspection, but there is active development of methods to measure non-visible changes [144]. There are multiple arthroscopic grading systems particularly for cartilage [145, 146] but their reliability due to subjective interpretation and user agreement is controversial [147]. Nonetheless it is often considered a gold standard to validate other radiological and clinical measures. Arthroscopy represents an invasive surgical procedure with associated risks, and its therapeutic benefits for OA are limited. Therefore it is not a routine assessment for OA.

1.2.9 Symptom - Structure Discord

There is considerable discordance between patients' clinical symptoms and structural joint tissue changes. This is particularly so when using radiographic assessments [148, 149, 150]. On average, nearly half of patients with symptoms of OA have no radiographic features and vice versa [150]. There is a general trend between symptoms and structural change in OA. However, at an individual level, severity of radiographic OA can bear little relation to the severity of the pain and disability experienced [151]. Different radiographic features have variable relationships to symptoms such as pain [reviewed in 152].

Radiographs are an incomplete and insensitive measure of joint pathology, and pain is a highly variable and multifactorial phenomenon. A meaningful relationship between the two on cross-sectional analysis is probably unlikely. A better association may exist between increase in radiographic severity and change in symptoms [153]. Other

potential factors contributing to the discord include technical aspects of radiograph acquisition, and criteria and grading of radiographic OA.

MRI has suggested that pathological changes not visualised by radiographs account for symptoms. Associations between symptoms and MRI synovitis, bone marrow lesions and subchondral bone attrition have been demonstrated [86, 85, 84, 152, 154]. The relative contributions of these tissue in the pain experience is unclear. MRI cartilage volume, however, does not seem to relate greatly with symptoms [155][reviewed in 152, 134]. Furthermore, a recent report states that "MRI does not improve the link between joint degeneration and the clinical expression of knee OA" and so the structure-symptom discord may be to some extent real [156].

1.2.10 Phenotypes

OA is not a single disease entity, but represents a final common pathway of joint failure. It encompasses an extremely heterogenous spectrum of risk factors and aetiologies. Furthermore, people with OA have equally heterogenous manifestations (pathological and clinical), progression and response to treatment. With this realisation, there is now a concerted effort to delineate well-characterised patient subgroups, or "phenotypes", that will promote our biological understanding, guide treatments and develop novel therapies [157, 158, 159, 160, 161] .

It is likely there will be many overlapping and evolving phenotypes based on:

- *Aetiology*: Risk factors and known causes [19, 162].
- *Location*: Generalised or single joint involvement, and particular joint(s) involved.
- *Structure*: Pathophysiology and pattern of tissue involvement within the joint [163].
- *Symptoms*: Pain experience, functional disability and their influencing factors [164].
- *Natural history*: Rate of progression and its nature i.e. inertia [165], phasic or continuous.

1.2.11 A New Staging Paradigm

Staging is an assessment of disease severity that is used to inform prognosis, guide treatments and evaluate their efficacy [166]. Severity can be defined in terms of symptoms and functional impairment, or pathological and structural changes. There are numerous methods and grading systems for assessing symptoms and structure in OA, and there is potential discordance between them. Consequently, there is no clear consensus on how best to stage OA. In 1949, Collins insightfully commented on his own macroscopic cartilage grading system that, "this grading system is, of course, arbitrary and in no way represents fixed stages in the progress of this disease" [167]. This statement still holds true for most of the current OA staging and grading systems. Furthermore, staging is likely to be specific for different OA phenotypes.

There is particular interest in identifying patients with "early OA", especially those at risk of progression, who may benefit from disease-modifying treatments. However, early OA is a frequently used term with variable descriptions and definitions in the literature, based on clinical, radiological or histological parameters. There is fortunately now a growing effort to standardise a definition for early OA [168, 169].

An alternative staging system to guide possible response to disease modification may be defined by crucial points in OA pathophysiology [Figure 1.5]. *Initiation* of joint failure occurs when there is imbalance in joint anabolism and catabolism. Joint failure progresses towards *end-stage OA*, defined as clinically severe and structurally advanced disease, refractory to non-operative management, suitable for joint replacement surgery. This usually involves significant cartilage loss with areas of bone-on-bone contact. *Joint replacement* effectively signifies joint death and organ transplant without which the patient continues in a state of end-stage OA. Prior to initiation, high risk individuals are amenable to preventative measures. Between initiation and the *point-of-no-return*, it is possible to reverse progression and restore a normal joint with regenerative treatments (i.e. reversible OA). Beyond the point-of-no-return, disease-modifying treatments can halt progression (i.e. modifiable OA). After the *event-horizon* there is inescapable progression to bone-on-bone end-stage OA (i.e. resistant OA), although it may be possible to slow this down by mechanical interventions such weight loss or osteotomy.

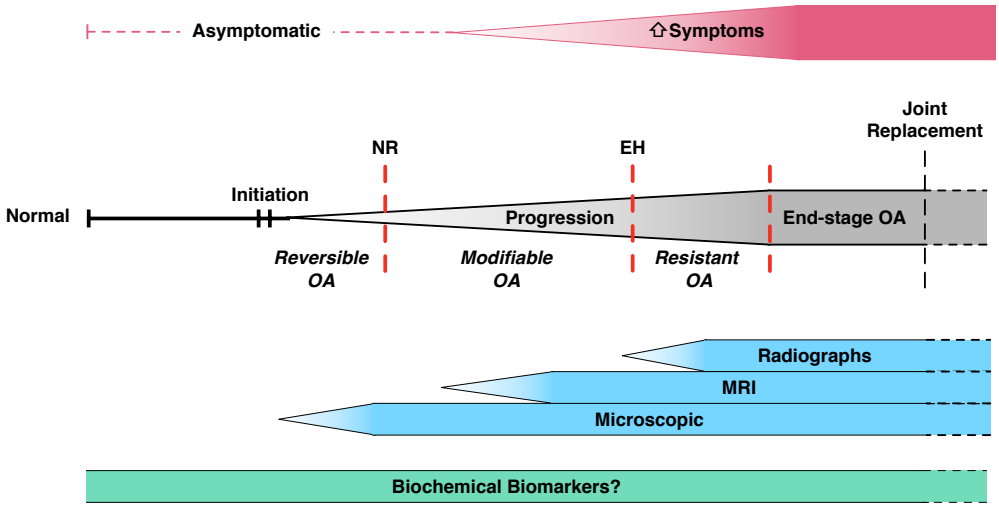


Figure 1.5 – A new staging paradigm for osteoarthritis. A theorised schematic is provided for the OA continuum from normal joint to end-stage knee OA and joint replacement. This involves initiation and progression of joint failure. Clinical symptoms are characterised by a silent asymptomatic phase, despite ongoing joint failure. Detection of pathological and structural changes depends on the method of assessment. For example, radiographic features represent relatively advanced phases of OA. Before the *point of no return* (NR), regenerative treatment can restore normal joint homeostasis and function (Reversible OA). After the *event horizon* (EH), there is unrelenting progression to end-stage OA although treatment may slow this down (Resistant OA). Between these two points disease-modifying interventions can stabilise the joint, by slowing or halting progression (Modifiable OA). It is possible that biomarkers may be able to identify all stages of the OA continuum.

It is possible that one or more of initiation, point-of-no-return and event-horizon may coincide.

1.2.12 Treatment

The treatment of OA aims to alleviate pain, improve function and quality of life, and ideally prevent progression. There is no cure.

Management strategies are categorised into non-pharmacological, pharmacological and surgical. Non-pharmacological treatments for knee OA include life-style and activity modification, patient education, weight-loss, physiotherapy, acupuncture, walking aids and orthotics and braces. There is evidence that weight-loss may have disease-modifying actions [170, 171]. Pharmacological treatments have centred on the analgesic ladder, topical analgesia and intra-articular injection of corticosteroid or hyaluronic acid viscosuplementation. There is an increasing use of nutraceuticals such as glucosamine and chondroitin sulphate, although their benefits are debatable [172, 173]. The role of arthroscopic treatments for knee OA continues to diminish

[174, 175], and the main surgical options are high tibial osteotomy [176, 177] and joint arthroplasty.

There are numerous evidence-based guidelines for the treatment of OA and most recommend a patient-centred and holistic multimodal approach [178, 179, 180, 181, 182]. However, for many patients joint replacement surgery is ultimately the only treatment option available. All guidelines advocate joint replacement for advanced OA in the presence of severe intrusive symptoms and functional limitations, associated with a poor quality of life, despite non-surgical management. Knee replacement has been established as a successful and cost-effective intervention particularly in older individuals [183, 184, 185]. Despite its efficacy, arthroplasty it is not without its risks, complications and drawbacks.

1.2.13 Lack of Disease-Modifying Therapies

Current treatment modalities for OA largely target the patient's symptoms and are largely palliative. This should not be viewed as a criticism. The most important thing to the patient is their pain and loss of function, and any strategy that alleviates these should be applauded. However, despite intense interest from academic institutions and the pharmaceutical industry, there are currently no safe and effective therapies that can preserve joint structure[1], modify progression or reverse the process of joint failure. The term therapy is used so as to be inclusive of non-pharmacological, pharmacological and potentially surgical interventions with disease modifying effect. Excellent reviews on the current state of disease modifying OA drugs (DMOADs) can be found in the literature [186, 187, 188, 189].

So why are there still no disease-modifying therapies? There are broadly speaking two major obstacles to the development DMOADs: the complexity of OA and the lack of adequate methods to measure it.

OA represents synovial joint failure and is the culmination of a heterogenous spectrum of risk factors and aetiologies. Its initiation, pathogenesis, and progression are all poorly

[1] High tibial osteotomy may be considered a disease modifying intervention in that it can delay joint replacement for on average 6 years in patients with medial compartment knee OA [176, 177].

understood. In particular the “point of no return” remains a mystery. Consequently, there may never be a panacea for OA. This multifactorial disorder will most likely require a multi-pronged attack. The drive to identify and characterise OA phenotypes will promote better pathophysiological understanding, diagnosis, monitoring and allow targeted phenotype-specific DMOAD therapy. Finally, it must not be forgotten that OA is still a mechanically driven process, albeit one that is biologically mediated [34, 190]. Therefore, expecting that DMOADs can be effective without correcting the abnormal biomechanical environment may be misguided.

There is no clear consensus on what constitutes clinically meaningful endpoints in trails of disease-modifying interventions in OA: patients’ symptoms or structural (pathological) change? Ideally it should be both [191]. However, available methods of assessing either are often unsatisfactory. Current regulatory body guidelines recommend that disease-modifying interventions require evidence of impact on both radiographic joint space width (JSW) and symptoms [192, 193]. As discussed previously, JSW is an indirect, inaccurate, unreliable and very lethargic measure of cartilage thickness that is unrepresentative of whole joint pathology. Radiographic features are indicative of advanced OA that may be refractory to disease modification. Furthermore, there can be considerable discordance between radiographic findings and the patients symptoms. Unfortunately, radiographs are the present “gold standard”. There is therefore a critical need for better methods of measuring OA to facilitate the development of disease modifying therapies.

1.3 Biomarkers

The term “biomarker” is a portmanteau of “biological” and “marker”. The formal definition established by the National Institute of Health (NIH) Biomarker Definitions Working Group is:

> **Biomarker**: A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [194].

This is a broad definition that encompasses physical examination findings, laboratory tests and radiological studies. Therefore, in reality, the use of biomarkers in clinical practice dates back to antiquity [195]. Early biomarkers were often identified as byproducts of clinical practice, based on biological insight into underlying physiology, and evolved with increasing scientific knowledge and technological progress. However, the current machine of biomarker discovery is founded on the drive to develop and assess novel therapeutic strategies, and uses a formidable armoury of biotechnologies. This is reflected in an alternative and more detailed definition of a biomarker by Lohmander and Eyre.

> A biomarker is a structural or physical measure, or cellular, molecular, or genetic change or feature in a biologic process that can be identified and monitored, with resulting diagnostic or prognostic utility. Biomarkers must be reliably and reproducibly measurable by standardised published methods, be used in several laboratories, and have undergone validation that they measure the intended process with sufficient specificity [196, 197].

Today, the term biomarker is associated with increasingly sophisticated methods and analytical tools to indicate disease trait (risk factors and mechanisms), disease state (diagnosis and staging) and disease rate (progression and therapeutic response) [198].

1.3.1 Need for Osteoarthritis Biomarkers

At the core of OA biomarker research is the quest for disease modifying therapies. Biomarkers may allow stratification of patients, in particular identifying those with, or at risk of, rapid structural (and/or clinical) progression. The annual rate of radiographic progression in unselected knee OA patients is very low and up to half of patients show no changes at all [131]. Consequently, under current regulatory requirements, DMOAD trials need to be very large and long if they are to have any hope of demonstrating a therapeutic effect beyond placebo. Such trials would be expensive and have trouble recruiting and retaining participants. Identification of "rapid progressors" would facilitate shorter, smaller (more powerful) and potentially cheaper trials linked to existing clinical and structural outcomes. Furthermore, this patient population is likely to benefit greatly from disease-modifying treatment in everyday clinical practice. Ideally, biomarkers would eventually supersede radiographs in defining clinical outcomes as well as being able to measure DMOAD response and guide dose titration.

Most structural joint changes and possibly symptoms represent OA beyond rescue. The search is on for biomarkers that can identify high-risk individuals or patients with early, and therefore, potentially reversible disease. Such biomarkers would reveal the ideal target population for preventative and early intervention for both DMOAD trials and clinical practice.

Beyond the immediate needs of the pharmaceutical industry, biomarkers will enrich our understanding of the pathophysiology of OA at a tissue-specific and whole joint level. Due to the limitations previously discussed, biomarkers that can replace radiographic parameters to diagnose OA and monitor its progress will have significant clinical utility. They may provide better correlation with symptoms, and allow better patient characterisation ("phenotyping") and staging to guide management with current treatment modalities.

1.3.2 Classification and Terminology

Biomarker research is associated with its own extensive vocabulary.

Biomarkers are detected by a diverse range of analytical tools and they can be categorised into 2 major groups according to the technology used to generate them [199].

- *"Wet Biomarkers"* correspond to genetic (DNA and RNA) and biochemical (peptides, proteins, carbohydrates, metabolites etc.) features and changes. Measurements are usually made in a bodily fluid such as blood (serum or plasma), urine or SF using numerous possible biotechnologies. Wet biomarkers are also referred to as soluble or biochemical biomarkers.
- *"Dry Biomarkers"* include structural parameters using imaging modalities (i.e. MRI), or symptom and functional assessments such as clinician and patient-reported scores, performed tasks, gait analysis and muscle strength testing.

Biomarkers, and in particular wet biomarkers, can also be considered as either "direct" or "indirect". Direct biomarkers identify known molecular processes within a given tissue (i.e. cartilage, bone and synovium), and indirect biomarkers reflect less clearly definable general joint change that may represent contributions from several events and tissues.

The Osteoarthritis Biomarkers Network have proposed the BIPEDS classification system aimed to improve the development and analysis of OA biomarkers, facilitate therapeutic trials and promote communication using a common lingo [Table 1.6] [200, 199]. Any particular biomarker may meet the requirements of more than one category in the BIPEDS system.

Biomarkers must be contrasted from clinical endpoints, which are ultimately the most important and credible parameters in assessing the benefits and risks of any management strategy or therapy.

- *Clinical endpoints* are characteristics or variables that reflect how a patient feels, functions or survives [194]. They include the *ultimate endpoint* that refers to survival or rate of other serious and irreversible morbid events, and *intermediate endpoints* that are symptom or function endpoints before the ultimate

Biomarker Type	Definition
Burden of Disease	Indicates the extent or severity of disease within a joint, or in terms of number of joints involved
Investigative	Insufficient evidence to be assigned one of the other biomarker categories
Prognostic	Predicts progression in patients with OA, or future onset in those with no OA at baseline
Efficacy of Intervention	Indicate efficacy of treatment in OA patients, or those at high risk for development of OA
Diagnostic	Differentiates patients with OA (or a OA subtype) from those not having OA. Identification of patients with a high risk for OA
Safety	Monitors for local or systemic adverse effects and can be used for dose titration

Table 1.6 – BIPEDS classification of osteoarthritis biomarkers. Adapted from Bauer et al. and Kraus et al. [200, 199].

endpoint [201]. In the context of knee OA, the ultimate endpoint would be joint replacement surgery and examples of intermediate endpoints could include onset of night pain or use of a walking aid. In general, the progression of OA is slow and the use of clinical endpoints, especially the ultimate endpoint of arthroplasty, to assess interventions may not be feasible. Consequently, one of the holy grails of biomarker research is the discovery of a surrogate endpoint.

- A *surrogate endpoint* is a measurement or biomarker that can substitute for a clinically meaningful outcome or endpoint, and therefore predict the effect of an intervention [194, 196]. An example of a surrogate endpoint is the use of viral load and CD4 cell counts as substitutes for opportunistic infection, AIDS and death in the evaluation of anti-retroviral dugs in patients with HIV infection. For knee OA, the search for the "virtual joint replacement" is slowly gaining momentum [94, 202, 203, 204]. Very few biomarkers will ever achieve the privileged status of surrogate endpoint, not in the least because it requires considerable and unequivocal supporting evidence.

Finally, one must distinguish between the concepts of biomarker validation and qualification [199].

- *Validation* refers to the process of verifying the analytical performance (i.e. accuracy, precision etc.) of a biomarker measurement or assay technique. It also includes consideration of pre-analytical (i.e. patient factors and sampling) and post-analytical (i.e. data management and statistical analysis) sources of error and variability.
- *Qualification* usually refers to the evidentiary process of linking a biomarker to a biological process, structural outcome or clinical outcome. Qualification against a clinical endpoint leads to surrogate endpoint status, but is a highly regulated and stringent process.

1.3.3 Biomarkers Sources

Dry biomarkers are almost synonymous with imaging investigations, and in particular MRI. This has been briefly discussed previously (Section 1.2.8.2) and further discussion of MRI biomarkers is beyond the scope of this work.

Biochemical biomarkers can be measured in tissue or biological fluids such as blood (plasma or serum), urine and SF. Joint tissues offer highly specific biomarkers at undiluted concentrations that more closely reflect their production and are less influenced by clearance from the joint. It may be possible to guide biopsies to areas of greatest or least structural pathology as desired. However, tissue biomarkers are rarely used in OA biomarker research. Biopsies are highly invasive procedures requiring at least local if not general anaesthetic. They are therefore unacceptable to patients and repeat measurements are impractical. Furthermore, the biopsy process induces injury that may compound the OA process and confound its future measurement.

Biofluids are far easier to obtain. Urine collection is completely non-invasive and is often part of routine clinical examination. Serum (and plasma) samples require venepuncture that is regularly conducted for basic laboratory investigations, and is tolerated by virtually all individuals. Arthrocentesis for SF is more invasive and may

require local anaesthesia in some instances. There is also a very low risk of introducing infection, but the procedure remains acceptable to many individuals.

1.3.4 Wet Biomarker Kinetics

The relationship between joint biology and biomarker measurements in various biological fluids is complex [Figure 1.6] [205]. Concentration is a static measure. It depends on both the production, which is usually of most interest, and clearance from a fluid compartment. Furthermore, production is influenced by the existing volume and activity of tissue releasing the marker. In the case of matrix components, there are added layers of complexity in that a marker may reflect matrix synthesis, degradation or both. Markers of degradation can originate from mature ("old") matrix, newly synthesised matrix and newly synthesised macromolecules before matrix assembly.

There are a number of (pre-analytical) factors that can affect biomarker levels including age, gender, BMI, diurnal rhythms, food intake, physical activity and comorbidities. These have been reviewed elsewhere [206, 207].

1.3.5 Which Biological Fluid?

1.3.5.1 Peripheral Fluids

Serum and urine biomarker levels reflect systemic skeletal tissue turnover. Although easy to collect, they are not necessarily specific to any particular joint or even to OA. Concentrations are influenced by other synovial joints (with or without OA) including facet joints of the spine and other skeletal tissues such as bone and intervertebral discs. Potentially other tissues, such as the bronchial tree that contains hyaline cartilage, may also confound measurements [209]. Furthermore, biomarker levels are subject to metabolism and clearance by the liver and kidney, and are therefore dependant on the health and functional capacity of these organs. In particular, urine biomarkers need to be corrected with creatinine excretion measurements. Due to these considerations, serum and urine biomarkers are more vulnerable to pre-analytical sources of variability

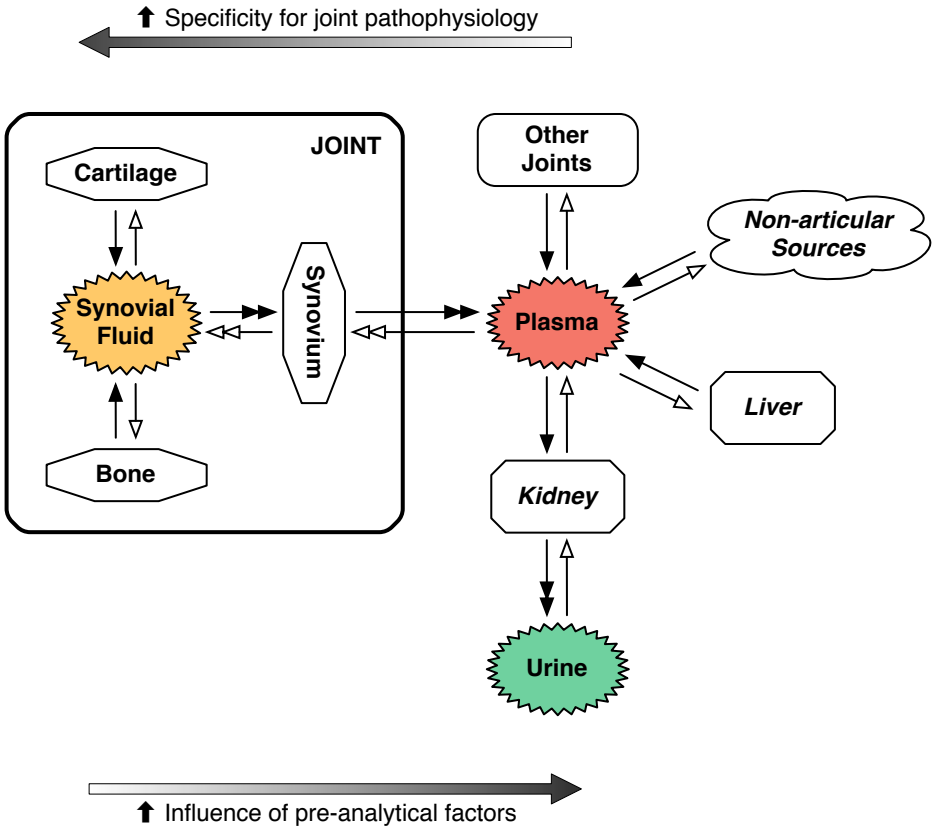


Figure 1.6 – Biomarker kinetics. A marker released from one of the 3 primary joint tissues depends on the biological process and the volume of tissue. It is diluted in the synovial fluid and may be masked from identification by hyaluronic acid macromolecules. The marker may be metabolised by the synovium, or indeed cartilage and exposed bone. It is cleared by the synovial lymphatics and capillary network, influenced by its size and charge (blood-joint barrier) and also degree of synovial inflammation. On entering the circulation, it undergoes further considerable dilution. Further contributions of the marker may be added from other synovial joints, and both skeletal tissues and non-skeletal tissues. It undergoes metabolism and clearance predominantly by the liver and kidney, but possibly other tissues. What remains can be filtered by the renal glomeruli to enter the urine, where levels may either be diluted or concentrated. Marker levels in any biofluid compartment can be affected by factors such as age, gender, BMI, diurnal changes, food intake and activity levels. However, these effects generally increase with "distance" from the joint. Overall, the specificity of a marker and most likely its sensitivity and responsiveness to biological processes of interest increase with "proximity" to the joint tissues. Adapted from Simkin & Bassett [208].

between and within the same subject, independent of OA status. Serum sampling may also be affected by the site of venepuncture [210]. Serum and urine biomarkers, therefore require considerable standardisation and validation before any attempts at qualification. Overall, these peripheral biomarkers are arguably more suited to generalised and poly-articular phenotypes of OA.

1.3.5.2 Synovial Fluid

SF is undoubtedly the most direct and specific measure of joint biology and pathophysiology. There are many examples of this in the literature. For example, Xu et al. recently demonstrated that SF, but not serum, levels of the chemokine CXCL12 was significantly correlated to radiographic knee OA severity [211]. This specificity is not exclusive to the knee or OA. SF concentrations of receptor activator of NF-$\varkappa$B ligand (RANKL) show significant correlation to clinical and radiographic parameters of severity in both hand OA patients and rheumatoid arthritis patients [212]. There were no significant associations for serum RANKL levels, particularly in the OA group. Incidentally, serum levels were able to distinguish both groups from controls. Given the poly-articular nature of hand OA and rheumatoid arthritis this study again confirms that peripheral markers are better suited for global pathology.

The concentrations of many popular candidate biomarkers, such as COMP and aggrecan neoepitopes to name two, have long been known to be significantly higher in SF than serum [213, 214]. Sometimes the superiority of SF measurements is not explicitly stated and hidden in the detail. A recent study measuring nitrotyrosine (NT), as a marker of oxidative stress, found SF and plasma levels to be significantly higher in OA than controls [215]. However, SF NT levels were double that in control, whereas plasma levels were ~40% greater and there was no significant difference in urine levels. For many potential biomarkers, the association between SF and peripheral concentrations are at best weak [216, and C. Jayadev, unpublished data].

There are practical advantages to these considerations. Higher concentration, increased specificity and relatively low systemic pre-analytical variability means that promising SF biomarkers can be demonstrated in relatively small study populations [217, 218]. In

particular, Giovanni et al. demonstrated that SF IL-6, macrophage chemotactic protein (MCP)-1 and fibronectin-aggrecan complex each significantly increased with worsening degree of ankle chondrosis, in an adequately powered study of just 36 patients [217]. High local concentrations may allow precious SF samples to be diluted, effectively increasing the volume available for multiple analysis.

The major challenges to SF biomarker analysis are that sample collection is more invasive and available volumes are lower than for peripheral fluids. SF volumes in OA joints are frequently low even in large joint such as the knee. There may be insufficient sample for the intended measurements, or may even result in a "dry tap". This is often a problem when attempting to obtain samples from normal or early OA individuals. Ultrasound guided aspiration can sometimes overcome this challenge. Some authors recommend joint lavage, but this introduces an uncontrolled dilution factor that needs to be formally corrected [219, 220]. Proponents of peripheral sampling will argue that longitudinal SF sampling as required for clinical trials is not feasible, but repeat aspirations have been been successfully conducted in the past [221, 222].

The SF compartment is not completely immune to pre-analytical considerations and there remain unanswered questions. Given the complex rheological properties of SF and variable and patchy synovitis of OA it is not known whether the site of sampling would affect biomarker measurements, particularly in a large and complex joint such as the knee. SF most likely represents the most dynamic source biomarkers due to its close proximity to the pathophysiology of OA. There is currently no information on the stability of SF measurements over short time periods during which there has been no structural or symptomatic change. Both these issues would have practical implications for the use and validation of SF biomarkers.

1.3.6 Observations on Current Trends

The literature for OA biomarker research is extensive and growing [Figure 1.7]. The current focus is firmly on biochemical biomarkers with <5% studies involving MRI biomarkers, although these are now gathering pace. It is not within the scope of this introductory chapter to provide an exhaustive review of this vast literature, which will

be selectively encountered in the proceeding chapters of the thesis. For now, readers are directed to any one of several excellent recent review articles summarising biochemical biomarkers in OA[2]. As well as the BIPEDS classification, potential biomarkers are grouped according to tissue of origin and pathophysiological processes [reviewed in 233]. Markers of cartilage synthesis and degradation, and markers of bone and synovial metabolism have received the greatest attention. These and others popular biochemical biomarkers under investigation are shown in Table 1.7.

There are some important general observations to be made. Despite the potential limitations discussed previously, the majority of publications report investigations into serum, plasma and/or urine. Just over 25% (562/2061) of studies involved SF analysis. A recent systematic review reports that urinary CTX-II and serum COMP have the best performance of commercially available urine and serum biomarkers [223]. However, the authors comment on variable study designs, heterogenous study populations, lack of consistent evidence for most commercially available biochemical markers and possible publication bias for less popular markers.

The workhorse for soluble marker measurements is the enzyme-linked immunosorbent assay (ELISA) and studies have traditionally measured one or a handful of molecules. However, there is an increasing use of ELISA-based multiplex immunoassay platforms such as Luminex and MesoScale Discovery. Wide-spectrum high-throughput methods such as proteomics and metabolomics show great potential and are being actively encouraged [reviewed in 234, 235]. Although their popularity is growing, they are not yet commonly used.

1.3.7 Whither Osteoarthritis Biomarkers?[3]

There are currently no reliable OA biomarkers. Despite extensive research, no biomarkers have been sufficiently validated or qualified to aid diagnosis, inform prognosis, or monitor progression or response to therapeutic intervention. None come close to representing a surrogate endpoints.

[2] References: 223, 224, 225, 226, 227, 228, 229, 230, 231, 232

[3] Section heading borrowed from Felson and Lohmander [236]

Biomarkers related to collagen metabolism
▶ C-terminal telopeptide of collagen type II (CTX-II)
▶ Type II collagen α chains collagenase neoepitope (α-CTX-II)
▶ Type II collagen propeptides (PIINP, PIIANP, PIIBNP, PIICP, CPII)
▶ Pyridinoline and Glc-Gal-PYD
▶ Type II collagen cleavage product (C2C)
Biomarkers related to aggrecan metabolism
▶ Core protein fragments (aggrecan neoepitopes, ARGS and FFGV fragments)
▶ Chondroitin sulfate epitope 846 and monoclonal antibody 3B3(–)
▶ Keratan sulfate
Biomarkers related to other non-collagenous proteins
▶ Cartilage oligomeric matrix proteins (COMP & deamidated, D-COMP)
▶ Fibulin (peptides of fibulin 3, Fib3-1, Fib3-2)
▶ Follistatin-like protein 1 (FSTL-1)
▶ Hyaluronan (hyaluronic acid)
▶ Matrix metalloproteinases (MMP-1, MMP-3, MMP-9, MMP-13 & TIMPs)
▶ YKL-40 (cartilage glycoprotein 39)
▶ Soluble receptor for advanced glycation endproducts (sRAGE)
Biomarkers related to other processes
▶ Inflammatory biomarkers: hs-CRP, IL-1β and IL-6 and COX-2
▶ Factors indicating fibrosis and complement proteins
▶ Adipokines (adiponectin, leptin, visfatin)
▶ Soluble receptor for leptin (sOB-Rb)
▶ Cellular interactions in bone (periostin)
▶ Wnt inhibitors (DKKs and SOST)
▶ Uric acid

Table 1.7 – Selected biomarkers currently being evaluated for osteoarthritis. COMP, cartilage oligomeric protein; COX-2, cyclo-oxygenase-2; Glc-Gal-PYD, glucosyl–galactosyl–pyridinoline; hs-CRP, high sensitivity C reactive protein; IL, interleukin; MMP, matrix metalloproteinase; PIIANP, N-propeptide IIA of type II collagen; PIIBNP, N-propeptide IIB of type II collagen; PIICP, C-propeptide of collagen type II; PIINP, N-propeptide II of type II collagen; SOST, sclerostin; TIMP, tissue inhibitor of matrix metalloproteinase. Reproduced from Lotz et al. [232].

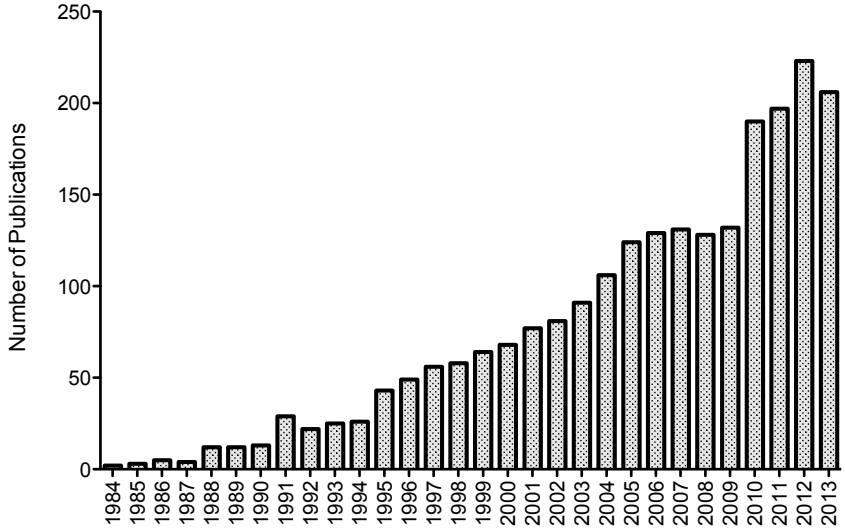


Figure 1.7 – Number of osteoarthritis biomarker publications per year. The PubMed bibliographic database was interrogated using the Advance Search Builder function. The search query was: "osteoarthritis" AND ["biomarker(s)" OR biochemical marker(s)"]. A total of 2319 publications were returned, of which 257 were review articles. Adding "synovial fluid" to the original query returned 610 (562 original) articles. Adding "MRI" to the original query returned 104 (89 original) articles. Adding "proteomic(s) OR "metabolomic(s)" to the original query returned 73 (53 original) articles. The searches were performed with an end-date of 30-Jan-2014. However, publications/year are presented until end of year 2013.

So why has the OA biomarker discovery machine failed? Ironically, the lack of disease-modifying therapies has hindered the ability to identify and assess the credibility of candidate OA biomarkers. This has been further compounded by the complexity and heterogeneity of OA, which biomarkers were supposed to help decipher. This scenario defines a self-fulfilling "terrible triad" in OA research. There is also no consensus regarding the choice of gold standard for biomarker qualification, structure or symptoms, and how to best measure them.

The synovial joint is a complex organ. Given the involvement of all its tissues and multiple pathophysiological pathways, it seems extremely unlikely that any single biochemical or imaging parameter will be able to succeed as an all-encompassing OA biomarker. Traditionally, studies have investigated single or only a small number of markers. Furthermore, there is a continued emphasis on serum and urine biomarkers. These are global measures, inherently lacking specificity for any particular joint and are severely vulnerable to systemic and pre-analytical variability. Relatively few studies attempt to control for such factors. The lack of specificity is compounded further

by study designs that only use a single non-OA control (usually “normal”) cohort. Differences in markers of matrix (cartilage or bone) turnover, matrix enzymes and soluble mediators such as inflammatory cytokines, growth factors etc. should not be considered as specific OA biomarkers because they may be present in other joint pathologies i.e. inflammatory arthritis. Most investigations target a single joint system, most frequently the knee, but these patients represent a heterogenous group. All things considered, most studies of potential urine or serum biomarker are therefore too underpowered to confirm their legitimacy.

Potential sources of technical errors and inconsistency between studies have also impeded biomarker discovery. These include factors such sample collection, storage and preparation; assay formats, reagents and manufacturers; quality control of assay performance parameters; and handling of raw data and statistical approaches. Due consideration to data handling and statistical analysis are often ignored. For example, the (mis)handling of censored and missing data, and inappropriate multiple comparisons can introduce considerable bias leading to erroneous conclusions [237].

1.3.8 The Road Forward

Prominent scholars, academic institutions and societies, the pharmaceutical industry and regulatory bodies have all acknowledged the lack of OA biomarkers success. There have been many recommendations and initiatives to rectify the situation [236, 199, 232, 231]

1.3.8.1 Multiple and Combination Biomarkers

There is unanimous agreement that multiple markers in combination will have the best hope of adequately reflecting the multiple domains of joint biology [238]. Wide-spectrum analysis of markers in serum and urine [239, 240] and SF [241] suggested the possibility of clusters of biomarkers related to specific pathogenic processes and tissues. From a statistical perspective, biomarker relationships and ratios are better able to differentiate OA disease stages [242] and combinatorial biomarkers have more power for statistically significant disease correlations [243, 244].

Newer proteomic technologies are promising and being actively encouraged [245, 246]. However, systems biology approaches are extremely expensive, require considerable bioinformatics support and are therefore not accessible to most laboratories. The development of existing promising biochemical biomarkers into multiplex assays will allow (time and cost) efficient analysis while minimising sample demands.

The combination of (multiple) biochemical markers with imaging, and in particular MRI, may have the greatest potential for all classes of biomarker [247]. An OARSI/Foundation for the National Institutes of Health longitudinal study is currently assessing a combination of 12 serum and urine biochemical markers with multiple MRI biomarkers in the large OAI cohort [248]. Although this study is likely to be hugely informative, it persists in ignoring the limitation of systemic (serum and urine) measures and makes no attempt to include SF biomarkers.

1.3.8.2 Standardisation and Statistics

Pre-analytical and analytical variability can be minimised through a process of standardisation to allow biomarker validation. The STARD Initiative provides a framework for the standards required [249]. Clinical assessments of structure (imaging) and symptoms need to be optimised and standardised to facilitate biomarker qualification. Furthermore, consensus is needed on statistical analysis, interpretation and reporting. Appropriate multivariate data analysis for multiple/combination biomarkers is especially important.

1.3.8.3 Phenotypes and Staging

The study of well-characterised phenotypes should facilitate biomarker discovery by overcoming some of the heterogeneity of OA. The spectrum of joint failure from initiation to joint replacement is broad and most likely phenotype-specific. Better definitions and methods of assessment are required to facilitate a useful staging system. In particular those for “early OA”. Although the emphasis seems to be on such “early” and pre-radiographic OA, biomarkers are needed for all stages of disease [199]. Further

studies are required on biomarker changes in relation to definite clinical endpoints such as joint replacement i.e. a virtual joint replacement [232].

1.4 Summary

Osteoarthritis (OA) is a very complex and heterogenous disorder of a sophisticated organ system, the synovial joint. Existing methods to detect early disease, assess severity and monitor progression have largely been inadequate. Consequently, treatment of this common and often debilitating condition is largely palliative and many patients ultimately progress to "organ transplant" by way of joint replacement surgery. There are currently no disease-modifying therapies that can delay, let alone halt or reverse joint deterioration in OA. It is hoped the development of such therapies will be fuelled by the promise of better patient characterisation, accurate and evolving MRI evaluation and specific biochemical biomarkers.

The lack of disease-modifying intervention has been mirrored and coupled to a continuing failure in OA biomarker discovery. Evolving paradigms of OA's complexity have led academic experts and learned societies to correctly advocate the hunt for multiple or combination biomarkers. Much hope has been placed in wide-spectrum, high-throughput "-omic" technologies. Furthermore, the largest ever collaborative biomarker initiative was launched to assess a dozen serum and urine biochemical markers in combination with multiple MRI parameters.

There are still no reliable OA biomarkers. The authors of *Osteoarthritis and Cartilage's* "Year in Review" for 2013 reflect that,

> "Frankly, it might be questioned whether we are making the most adequate steps ahead and we may have to take a step back to reconsider our approaches." [231]

The authors go on to make a strong recommendation for SF biomarker discovery. Serum and urine have traditionally been the major sources for biochemical biomarkers despite their considerable inherent flaws. SF is undoubtedly the most direct and specific measure of joint biology. However, there remain challenges to SF analysis due to its complex viscous properties and there are unexplored practical considerations regarding sampling. Will these issues prevent the discovery of useful combination SF biomarkers?

1.5 Scope of the Thesis

The purpose of the thesis is to demonstrate the utility of wide-spectrum SF analysis for biomarker discovery in knee OA. This will ultimately be achieved by defining a diagnostic SF "fingerprint" for end-stage knee OA from a diverse range of soluble molecules with pathophysiological relevance. However, SF is a difficult material to study, posing challenges in terms of quantity and inherent quality. Therefore, a number of important, previously neglected, practical considerations of SF collection and analysis will be addressed during the journey of biomarker discovery.

This thesis focuses on SF as a source of combination biomarkers in knee OA. Although not a direct exploration of OA pathophysiology, it is hoped that this work will improve the understanding of OA. The thesis is not intended to be a comparison between SF analysis and other biological fluids i.e. serum and urine, or a discourse into statistical methods for biomarker research.

1.6 Detailed Objectives

1. **Optimise and validate methods for multiplex SF analysis**.

   SF is a difficult biological sample to analyse and the limited volumes available from OA knees means that it is usually impossible to measure multiple markers by ELISA. Despite the increasing use of multiplex technologies, they have not been optimised, validated or compared for SF analysis.

2. **Determine if the anatomical site of sampling affects SF measurements.**

   It is assumed that SF circulation and free diffusion ensure equilibrium of its composition amongst all three knee joint compartments. If incorrect, there are major implications on the practicalities and required level of consistency for SF sampling, particularly when there is localised joint involvement.

3. **Establish whether SF measurements are stable over short intervals**.

It is also not known if the SF biology remains constant for a given clinical and/or structural severity or is subject to short-term fluctuations that would invalidate its use for clinical decision making.

4. **Assess whether structurally progressive phenotypes of end-stage knee OA demonstrate differences in their SF biology.**

   Structural phenotypes in end-stage knee OA provide a potential framework to discover burden of disease markers since the "volumetric" extent of degenerative changes within the knee increases with the number of affected joint compartments. Furthermore, it is not known whether the high topographic consistency of structural phenotypes such as anteromedial knee OA is reflected in its SF biology.

5. **Define a "biological fingerprint" for end-stage knee OA from a diverse range of pathophysiologically relevant SF analytes**.

   Given the heterogeneity and pathological complexity of OA, a multi-marker approach is likely to be more successful than a search for a single biomarker. The discovery of a combination biomarker signature for end-stage knee OA in patients undergoing arthroplasty could represent a true surrogate endpoint. A biochemical "virtual knee replacement" could be a decision-making tool for clinical practice and development of disease-modifying therapies. A successful discovery and its approach would initiate and stimulate substantial further biomarker research.

1.7 Thesis Structure

The following chapter (Chapter 2) and its linked appendices outline the general methodology for the experimental work in the thesis, which is presented in two parts:

Part I (Chapters 3 to 6) is concerned with methods development and important practical aspects of SF collection and analysis for OA biomarker research.

- Chapter 3 explores how SF treatment with hyaluronidase, to reduce viscosity, affects the performance of multiplex immunoassay platforms such as Luminex and MesoScale Discovery (MSD).

- Chapter 4 provides a detailed analytical comparison between Luminex and MSD multiplex immunoassay platforms during SF analysis of patients with knee OA. This comprehensive work attempts to validate these platforms for SF analysis in OA biomarker research.

- Chapter 5 tests the assumption that the composition of SF is at equilibrium throughout the knee joint. The study determines whether the anatomical site of sampling is an important consideration affecting SF measurements when there are considerable regional differences in the extent of OA disease pathology as in anteromedial knee OA.

- Chapter 6 assesses the short-term temporal stability of SF measurements in patients with end-stage knee OA over a short time interval during which there has been no clinical or structural disease progression.

Part II (Chapters 7 to 9) applies the findings of Part I to explore the SF biology of end-stage knee osteoarthritis (OA) in terms of a wide spectrum of physiologically relevant cytokines, chemokines, growth factors, matrix enzymes and markers of cartilage turnover with a view to biomarker discovery.

- Chapter 7 details the study design, methodology and statistical approaches used to answer the research questions of Chapters 8 and 9.

- Chapter 8 investigates potential differences in the SF biology of well-characterised structural phenotypes of end-stage knee OA with increasing joint compartment involvement. This investigation assesses both the measured concentrations and variability of biological endpoints between the progressive structural phenotypes.

- Chapter 9 represents the culmination of the thesis and addresses the overarching aim to discover a "biological fingerprint" of end-stage knee OA. The study utilises samples from patients with knee injury and inflammatory arthritis as reference cohorts to maximise the specificity of any findings to end-stage knee OA.

Each experimental chapter addresses its own stated hypotheses and ends with a thorough discussion with reference to the relevant literature. Chapter 10 offers an

overall summary and discussion of the work in thesis, but given the primary aim of the thesis, concentrates on the finding of Part II. The final chapter, Chapter 11, makes recommendations for future work prompted by this thesis.

Chapter 2

General Methodology

2.1 Introduction

This chapter outlines the general methods used for the forthcoming studies within the thesis. Patient recruitment, assessments and sample collection are discussed in detail. Together with the related appendices, a description of immunoassays and their performance parameters for quality control and validation are also provided.

The first part of this thesis focuses on methods development and important practical aspects of synovial fluid (SF) collection and multiplex analysis for osteoarthritis (OA) biomarker research. The findings will be utilised for subsequent SF biomarker discovery in end-stage knee OA in the second part of the thesis. Additional methodology for each study is provided in the relevant chapters. Furthermore, Chapter 7 summarises the study design and methods specific for Part II of the thesis.

2.2 Ethics

All studies within this thesis were granted both regional and institutional ethical approval, and all participants gave their written consent. The studies were conducted in full compliance with the Human Tissue Act (2004) and Declaration of Helsinki (2013).

Patients of both sexes, above the age of 16 years, due to undergo knee surgery at the Nuffield Orthopaedic Centre, Oxford were invited to participate in the study. Patients undergoing arthroplasty for (end-stage) knee osteoarthritis (OA), and arthroscopic surgery for cruciate ligament and/or meniscal were recruited under the umbrella of the Oxford Musculoskeletal BioBank (OMB, REC Ref: 09/H0606/11). Additional ethical approval was obtained to collect control SF from the asymptomatic non-operative "good knee" of patients undergoing arthroscopic knee surgery (REC Ref: 10/H0721/19), and for longitudinal SF collection (REC Ref: 11/SC/0549). All patients gave their written informed consent.

Inflammatory SF samples from patients with rheumatoid or psoriatic arthritis were collected during a previous study (REC Ref: 06/Q1606/139), and were a gift (from Dr. Catherine Swales).

2.3 Patient Cohorts

Knee OA was diagnosed according to the American College of Rheumatology (ACR) clinical and radiological criteria [89]. End-stage knee OA was defined as clinically and radiologically advanced disease, refractory to non-operative management, being treated by knee replacement surgery.

Anteromedial gonarthrosis (AMG) [250]: Patients with degenerative changes confined to the antero-medial portion of the medial tibio-femoral articulation. The lateral tibio-femoral articular surfaces are well preserved with an intact meniscus. This confined pattern of disease was confirmed radiologically (radiographs and/or MRI) and intra-operatively. Furthermore, patients have intact and functioning anterior and posterior cruciate ligaments, fully correctable varus deformity and little ($<15°$) to no flexion deformity. Patients were categorised based on radiological and/or intra-operative evidence of patello-femoral OA changes as:

- Isolated anteromedial osteoarthritis (AMOA) or
- Anteromedial OA plus patello-femoral OA (AMOAp)

Lateral compartment osteoarthritis (LCOA): Patients with degenerative changes confined to the lateral tibio-femoral articulation. The central part of the lateral tibial articular surface and the posterior aspect of the femoral condyle are usually involved. The medial tibio-femoral articular surfaces are well preserved. This confined pattern of disease was confirmed radiologically (radiographs and/or MRI) and intra-operatively. Furthermore, flexion deformity is uncommon and valgus deformity is usually fully correctable.

Tri-compartmental osteoarthritis (TCOA): Patients with OA pathology affecting all three compartments of the knee joint: medial and lateral tibio-femoral, and patello-femoral articular surfaces. This global pattern of disease was confirmed radiographically and intra-operatively. Patients often have anterior cruciate ligament damage, fixed varus or valgus deformity and flexion deformity.

Knee injury (Injury): Patients with anterior cruciate ligament and/or meniscal injuries without any clinical, radiological (radiographic and MRI) or arthroscopic evidence of articular surface degeneration, osteochondral defects, subchondral remodelling or osteophytes, or significant synovitis. The median interval between injury and surgery was 6.5 months (interquartile range 4 – 9.75 months) [Appendix F.3].

Inflammatory arthritis: Patients with rheumatoid arthritis (RA) or psoriatic arthritis (PsA) affecting the knee joint. Patients were on a range of anti-inflammatory and/or disease modifying treatments [Appendix F.3].

Control: Patients with completely asymptomatic knees with normal clinical examination. Individuals were screened to exclude any potential history of knee pathology. Collection of normal control SF was attempted from the asymptomatic non-operative "good knee" of patients undergoing knee surgery for cruciate ligament or meniscal injuries.

Cohort	Description
AMOA	Patients with isolated anteromedial knee OA undergoing either primary unicondylar or total knee arthroplasty
AMOAp	Patients with anteromedial knee OA and patello-femoral OA undergoing either primary unicondylar or total knee arthroplasty
LCOA	Patients with lateral compartment knee OA undergoing either primary unicondylar or total knee arthroplasty
TCOA	Patients with tri-compartmental knee OA undergoing primary total knee arthroplasty
Injury	Patients undergoing surgery for cruciate or meniscal injuries without evidence of OA changes (confirmed by radiographs, MRI and at surgery)
Inflammatory	Patients with rheumatoid or psoriatic arthritis of the knee, with or without disease-modifying treatments
Control	Patients undergoing knee surgery, whose other non-operative "good knee" is asymptomatic, non-osteoarthritic and has not had previous surgery

Table 2.1 – Description of patient cohorts recruited for thesis experiments. The end-stage knee OA (AMOA, AMOAp, LCOA and TCOA) and Injury samples were prospectively collected. Inflammatory samples were retrospectively obtained as a gift and used as "positive controls". Note that collection of Control samples was unsuccessful.

2.4 Exclusion Criteria

Patients were excluded from prospective SF collection if there was a history of: rheumatoid or other inflammatory arthritis (in any joint); septic arthritis; post-traumatic OA; previous meniscal or cruciate injury with or without surgical intervention; previous arthroscopic surgery, high-tibial osteotomy or arthroplasty (partial or total) in index joint; systemic or intra-articular glucocorticoid within 6 weeks of sample collection; intra-articular viscosupplementation within 6 weeks of sample collection.

In addition to the above, patients were excluded from recruitment into the injury group if there was a history or clinical symptoms of OA in any joint system.

2.5 Patient Assessments

Clinical and structural (radiological and/or intra-operative) assessments were made for patients in the end-stage knee OA and injury cohorts. No assessments were available for patients with retrospectively obtained inflammatory SF samples.

2.5.1 Clinical Evaluation

Recruited patients' age and gender were recorded. Participants were asked to complete the Oxford Knee Score (OKS) questionnaire at pre-operative assessment, or on the morning of surgery if not previously obtained. The OKS is a 12-question patient reported outcome measure that assesses pain and function over the previous 4 weeks [112, 113]. Each question is scored on a Likert scale from 0 (worst) to 4 (best) giving a total score ranging from 0 (most severe symptoms) to 48 (asymptomatic). In addition to the traditional total score, the OKS was also split into "functional" and "pain" components as recently described [251]. In brief, the subtotal of the 7 pain-related questions is multiplied by 5; the subtotal of the 5 function-related questions is multiplied by 3.57. This gives separate pain and function indices from 0 (worst) to 100 (best).

2.5.2 Radiological Evaluation

Patients were assessed with standardised weight-bearing anterior-posterior and lateral views, and a patella skyline axial view of the index knee. These were conducted at pre-operative assessment clinic. Patients awaiting arthroscopic surgery, also underwent MRI scanning of the index knee within 3 months of surgery.

Radiographs were evaluated using the Kellgren and Lawrence (KL) and OARSI Atlas grading systems [Tables 2.2 and 2.3] [114, 119]. A KL grade was given to each knee joint compartment. The grade for the most degenerate joint compartment corresponded to the overall KL grade for the joint as a whole. A KL grade $\geq$2 for the patello-femoral (PF) compartment defined the presence of PFOA. This was confirmed intra-operatively by macroscopic assessment. By definition all patients in the TCOA cohort had PFOA in addition to medial and lateral compartment OA. OARSI grading was conducted with reference to the atlas provided in the original publication. In addition to the total OARSI score, patients were also given joint space narrowing (JSN) and osteophyte (OP) subtotal scores. The JSN score (0 – 6) was the sum of the medial and lateral tibio-femoral JSN scores. The OP score (0 – 12) was the sum of the medial tibial, medial femoral, lateral tibial and lateral femoral OP scores.

Grade	Definition	Description
0	None	No osteoarthritis
1	Doubtful	Doubtful narrowing of joint space and possible osteophytic lipping
2	Minimal	Definite osteophytes and/or possible narrowing of joint space
3	Moderate	Moderate multiple osteophytes, definite narrowing of joint space and some sclerosis and possible deformity of bone ends
4	Severe	Large osteophytes, marked narrowing of joint space, severe sclerosis and definite deformity of bone ends

Table 2.2 – Kellgren & Lawrence radiographic grading system

Radiographic Feature	Site	Severity
Osteophyte	Medial tibial plateau	0 - 3
	Medial femoral condyle	0 - 3
	Lateral tibial plateau	0 - 3
	Lateral femoral condyle	0 - 3
Joint Space	Medial compartment	0 - 3
Narrowing	Lateral compartment	0 - 3
Other	Medial tibial sclerosis	0 or 1
	Medial tibia attrition	0 or 1
	Lateral femoral sclerosis	0 or 1

Table 2.3 – OARSI Atlas radiographic grading system for the knee

MRI scans were assessed qualitatively for the presence of OA features such as cartilage thinning, synovitis, subchondral bone marrow lesions and cysts, and osteophytes. The primary use of MRI scans was to exclude OA in patients with knee injuries.

A single evaluator (Chethan Jayadev) conducted all radiographic assessments. The reliability of KL and OARSI grading was formally tested using randomly selected radiographs from the end-stage knee OA and injury cohorts [Appendix E]. The reliability was comparable to that generally reported in the literature and by the Osteoarthritis Initiative [252, 253, 254]. The MRI scans were reported by consultant musculoskeletal radiologists.

2.5.3 Intra-operative Assessments

Intra-operative assessments of medial, lateral and patello-femoral joint compartments were made for patients undergoing unicondylar and total knee arthroplasty to confirm structural phenotyping of end-stage knee OA. A diagnostic survey was performed in all patients undergoing arthroscopic surgery for anterior cruciate and/or meniscal injury to exclude osteochondral defects, OA changes to the articular surfaces, osteophytes and significant synovitis. Experienced fellowship-trained orthopaedic surgeons preformed all surgery and intra-operative assessments.

2.6 Synovial Fluid

2.6.1 Sample Collection and Storage

Synovial fluid (SF) samples from patients in the end-stage knee OA cohorts were obtained by needle aspiration after superficial dissection, but prior to arthrotomy to avoid contamination with blood. Unless otherwise stated, SF was obtained from the medial knee compartment for patients in the AMOA, AMOAp and TCOA cohorts and lateral compartment for patients with LCOA. For patients in the injury cohort, SF was needle aspirated from the patello-femoral compartment after routine skin preparation and extremity draping, but prior to any surgical incisions. Ultrasound

guided needle aspiration of the asymptomatic non-operative "good knee" of control patients was attempted under aseptic technique in the operating theatre (after induction of anaesthesia) prior to routine skin preparation and extremity draping of the operative knee. However, patient recruitment and collection of control samples was abandoned after 5 failed attempts. In all cases, lavage samples were not taken because of potentially variable and uncontrolled dilution that would make comparisons unreliable.

Samples were placed in sterile additive free specimen pots and stored immediately at 4°C pending processing within 4 hours. Before storage, SF was centrifuged at 3000g for 25 minutes at 4°C to separate solid debris and cells. The supernatant was then aliquoted into separate 500μl microfuge tubes, snap frozen in liquid nitrogen and stored at -80°C until analysis.

The collection of all SF samples was conducted or directly supervised by a single researcher (Chethan Jayadev), who also performed initial processing for storage of all SF samples.

2.6.2 Preparation for Analysis

SF aliquots were thawed to room temperature and clarified at 10000g for 10 minutes prior to use. The supernatant was carefully pipetted for analysis.

2.7 Immunoassays

SF samples were analysed using immunoassay methods such as enzyme-linked immunosorbent assays (ELISA) and its multiplex derivatives, Luminex xMAP and Meso Scale Discovery. These represent relatively cheap and accessible technologies in comparison to proteomic methods. A full description of immunoassays and these technologies can be found in standard texts and manufacturer literature [255, 256, 257]. A brief summary is provided in Appendix A.1 for information.

A full list of immunoassays, materials and equipment used in the thesis is provided in Appendix B. All commercially sourced immunoassays were conducted according

to the manufacturer's protocol. Prototype and custom MSD assays were performed using optimised in-house protocols following MSD guidelines and MSD recommended reagents.

2.8 Quality Control and Assay Performance

2.8.1 Standard Curves

Luminex, MSD and ELISA signal data were exported from native software to MasterPlex QT 2010 (Hitachi Solutions, San Francisco, USA) to standardise analysis. Background (blank) subtracted signals were used for calibrator points and samples. A 5-parameter nonlinear regression model with $1/y^2$ weighting was used for standard curve fitting and concentration calculation [258, 259]. Acceptance criteria for back-calculated concentrations of the standard curve were a relative error within $\pm$20% of nominal and a coefficient of variation of $\leq$20% for replicates [260, 261, 262, 263]. At least 75% of the calibrator points needed to meet these criteria for the standard curve and assay to be acceptable.

2.8.2 Assay Performance Parameters

Immunoassay performance parameters include:

- Bead events (Luminex only)
- Limit of detection (LOD): the lowest concentration that can be reliably distinguished from blank measurements and is therefore qualitatively present in the sample
- Lower and upper limits of quantification (LOQ and ULOQ): the lowest and highest concentration of analyte that can be measured with appropriate accuracy and reliability
- Intra-assay precision in terms of coefficient of variation (CV) for measured replicate concentrations

- Recovery as a measure of accuracy and matrix interference

A detailed description of these parameters in provided in Appendix A.2.

Average LODs, LOQs and ULOQs for each analyte [Appendix C] were determined from all assay plates that were run during the course of the thesis, and then applied to the analyte concentrations from patient samples analysed in the proceeding studies. The intra-assay precision was described by the coefficient of variation (CV) for measured replicate concentrations. A CV $\leq$20% was considered acceptable, unless replicate concentrations were between LOQ and 2LOQ, or $^3/_4$ ULOQ and ULOQ when $\leq$25% was satisfactory.

2.9 Statistics and Software

A broad range of statistical procedures was used to interrogate the data. The Shapiro-Wilks test was used to assess continuous data for a normal distribution and appropriate transformations were subsequently applied if required. The natural log (to the base e) was used for logarithmic transformations. Specific statistical tests will be outlined for each hypothesis or question as they are encountered in the following chapters. In all cases $p<0.05$ was considered significant. The family-wise error (FWE) from multiple comparisons was corrected using Holm's step-down procedure (Holm-Bonferroni method) unless otherwise stated [264]. An adjusted p-value (q-value) <0.05 was considered significant.

Unless otherwise stated, statistical analyses and graphing were performed using STATA version 12.1 (StataCorp, College Station, Texas, USA) and PRISM version 5.0 (GraphPad, San Diego, California, USA), respectively.

Part I

Methods Development and Practical Considerations

Chapter 3

Synovial Fluid Treatment To Improve Multiplex Immunoassay Performance

3.1 Introduction

Synovial fluid (SF) is a difficult material to study. It poses challenges to biomarker research in terms of quantity and inherent quality.

Inflammation, in terms of synovitis and effusion, in osteoarthritis (OA) is far less than in inflammatory arthritis. Therefore, osteoarthritic joints may offer insufficient SF volumes to allow analysis for multiple molecules in a single sample by conventional enzyme-linked immunosorbent assays (ELISA). Fortunately, multiplexing immunoassay technologies such as Luminex xMAP and MesoScale Discovery (MSD) can maximise the amount of biological information that can be gained from small volume samples. However, personal early experience with the polystyrene-bead Luminex platform for cytokine measurements in neat SF samples from knee OA patients was unsatisfactory. The number of bead events was mostly suboptimal, and measurements from technical replicates of the same SF sample displayed low and variable repeatability. Assay precision is essential for biomarker discovery, particularly when characterising and comparing patient cohorts [199]. One can potentially correct for measurement error (bias), but it is not possible to correct for imprecision.

SF is a non-Newtonian biological fluid with a complex non-homogenous matrix [11]. Furthermore, the viscosity of SF from OA joints is greater than from inflammatory joints [265, 266]. Methods to reduce SF viscosity by dilution and/or hyaluronidase treatment have been described before for a variety of analytic techniques including immunoassays [267, 268], two-dimensional gel electrophoresis [269] and automated differential cell counting [270, 271].

The work in this first experimental chapter aims to determine whether reducing SF viscosity improves intra-assay precision on the Luminex and MSD multiplex immunoassay platforms. In particular, SF treatment with hyaluronidase will be compared to dilution alone. The study is not intended to be a comparison of multiplex platforms, which is conducted in Chapter 4.

3.2 Null Hypotheses

1. The intra-assay precision of the Luminex and MSD platforms, as defined by the coefficient of variation for measured replicate concentrations, is not influenced by SF dilution or hyaluronidase treatment.

2. The intra-assay precision of either platform, determined by the coefficient of variation for measured replicate concentrations, is equivalent for SF samples treated by dilution or hyaluronidase.

3.3 Methodology

3.3.1 Synovial Fluid Treatments

SF samples were obtained from 16 patients undergoing primary knee replacement for end-stage knee OA.

All samples underwent identical processing. SF samples were thawed to room temperature (RT) and clarified at 10000g for 10 minutes. The supernatant of each sample was then treated by the following methods: no preparation (neat), 2-fold dilution (2xD), 4-fold dilution (4xD) or treatment with 2mg/ml hyaluronidase (HAse).

For each comparison experiment, the diluent provided by the manufacturer for calibrator standard preparation was used to make sample dilutions and prepare a 4mg/ml HAse solution. After dilution, samples were vortexed for 5 seconds. HAse treatment entailed 1:1 volume mixture of SF with 4mg/ml HAse, vortexing for 5 seconds and incubation at RT for 1 hour on a shaker. Samples were centrifuged at 1000g for 5 minutes and the supernatant used for the assay. The end result was 2-fold sample dilution with 2mg/ml (~1200U/ml) HAse digestion.

Neat versus HAse treatment was compared in all 16 SF samples using the polystyrene-bead Luminex (pLuminex) cytokine 10-plex. Each sample was prepared by both methods and run in triplicate, but neat and HAse treated aliquots were run on separate plates. All 4 methods (neat, 2xD, 4xD and HAse) were compared in 7 SF samples on the magnetic-bead Luminex (mLuminex) IL-6 plex and MSD (pro-inflammatory 4-plex) platforms. Each sample was prepared by all methods and run in triplicate on the same plate. All immunoassays were conducted according to the manufacturer's protocol. Figure 3.1 summarises the experiments conducted to compare SF treatment protocols.

3.3.2 Performance Parameters

The number of bead events per target (for Luminex assays only), intra-assay precision, cytokine concentration (corrected for dilution) and their recoveries were compared for each treatment protocol.

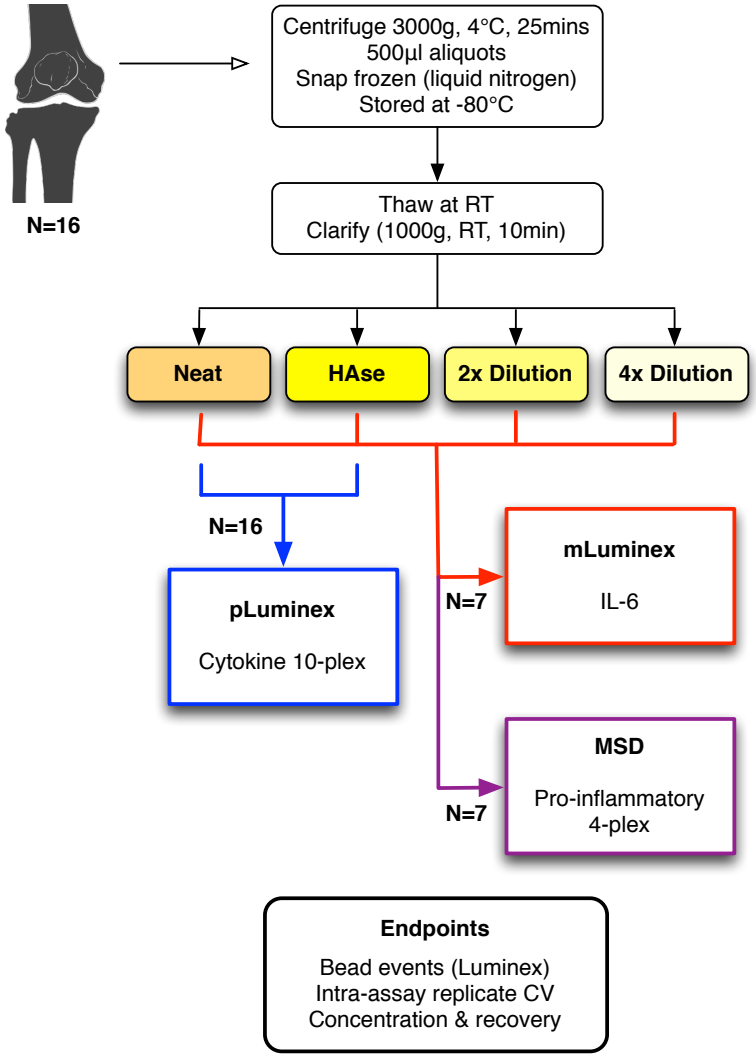


Figure 3.1 – Within-subject experimental design for comparison of SF treatments on the Luminex and MSD platforms. Sixteen SF samples were harvested and underwent identical processing. Each sample was split into aliquots for treatment by each method: untreated (neat), 2-fold diluted, 4-fold diluted and hyaluronidase (HAse) treated. Treatment methods were compared on a polystyrene bead Luminex (pLuminex) cytokine 10-plex (IL1-β, IL-6, TNF-α, IL-4, IL-8, IL-10, GM-CSF, IL-5, IFN-γ, IL-2); magnetic-bead Luminex (mLuminex) IL-6 single-plex, and a MesoScale Discovery (MSD) pro-inflammatory 4-plex (IL1-β, IL-6, IL-8 and TNF-α). The number of bead events per target (for Luminex assays only), intra-assay precision, cytokine concentration (corrected for dilution) and their recoveries were compared for each treatment protocol.

Bead Events

The number of bead events were categorised as "poor" (<35 counts), "good" (35 to 49 counts) and "excellent" ($\geq$50 counts). Absolute bead events were not used because the Luminex system continues to acquire beads until the minimum user-defined bead count is achieved in all wells for each bead region (target assay) or the system times out. The data from SF samples were defined as "functional" only if at least 2 replicates had $\geq$35 bead events/region; otherwise they were defined as "unusable". Only measured concentrations from "functional" samples were used for further analyses.

Intra-assay Precision

The intra-assay precision was determined by the coefficient of variation (CV) for measured replicate concentrations. A CV $\leq$20% was considered acceptable, unless replicate concentrations fell on the lower tail end of the standard curve when 25% was the threshold.

Dilution and Treatment Recovery

The percentage recovery of SF aliquots prepared by each method was calculated using the formula given in Appendix A.2. The mean replicate concentration calculated from the standard curve was used. The expected concentration for each sample was defined as the concentration for the neat aliquot, which therefore has a recovery of 100% and is not presented. Only samples above the LOQ were used for recovery analysis. Dilution recovery is given for 2-fold and 4-fold diluted aliquots, and treatment recovery for HAse treated samples. Dilution recovery was also used as a measure of assay parallelism. Mean recoveries and 95% confidence intervals (95%CI) are reported. Recoveries within $\pm$20% of expected were considered acceptable.

3.3.3 Statistical Methods

Matched (paired) and repeated measure statistical tests were used where possible since every SF sample was treated by each method under comparison i.e. within-

subject experimental design. Means with 95% confidence intervals (95%CI) are given for normally distributed data. Nonparametric data was natural log transformed and confirmed for normality. Back-transformed (geometric) means (GM) and their 95% confidence intervals (95%CI) are given for transformed data. 95%CIs were left censored at zero if needed as concentrations and CVs cannot be negative. Parametric statistical tests were then used throughout. $P<0.05$ was considered significant.

3.4 Results

3.4.1 Luminex Assays

The standard curve wells achieved the manufacturer's recommended bead count for all bead sets in each experiment. Standard curve points for all analytes in each of the experiments met the predefined acceptance criteria.

In the pLuminex 10-plex assays, HAse treatment significantly improved the number of wells with good or excellent bead events for each bead region [Figure 3.2A]. Overall, 95% of wells were good or excellent with HAse treatment compared to only 48% when untreated. There were significantly more SF samples with functional data for each assay after HAse treatment [Figure 3.2B]. Nearly all HAse treated samples (>95%) were functional for all bead regions compared to only 52% for untreated samples.

Only IL-6 and IL-8 were quantifiable in the 16 SF samples analysed. The intra-assay CV for functional samples was significantly lower after HAse treatment for both IL-6 (5.7% vs. 22.2%) and IL-8 (5.0% vs 28.3%) [Figure 3.3]. The dilution-adjusted concentration of both IL-6 and IL-8 was greater after HAse treatment compared to analysis of untreated neat samples, although this reached statistical significance only for IL-8 [Figure 3.4]. This was reflected in recoveries of >200% for both analytes after HAse treatment.

All plate wells in the mLuminex IL-6 assay achieved excellent (>50) bead events irrespective of the method of sample preparation, and therefore all samples were functional. HAse treatment significantly improved the intra-assay CV of measured IL-6 concentration (7.5%) compared to neat (28.2%), 2-fold diluted (26.8%) and 4-fold diluted (24.4%) [Figure 3.5A]. Dilution alone did not offer any significant improvement in CV compared to analysis of neat samples. Moreover, every HAse treated aliquot achieved a CV<20%. The dilution-adjusted concentration of IL-6 in the magnetic bead Luminex assay was greatest when samples were HAse treated ($p<0.001$) [Figure 3.5B]. There was no significant differences in corrected concentration between neat, 2-fold and 4-fold diluted samples. The recovery of HAse treated samples (>300%) was more than double 2-fold diluted and more than 1.5 times 4-fold diluted samples ($p=0.03$)

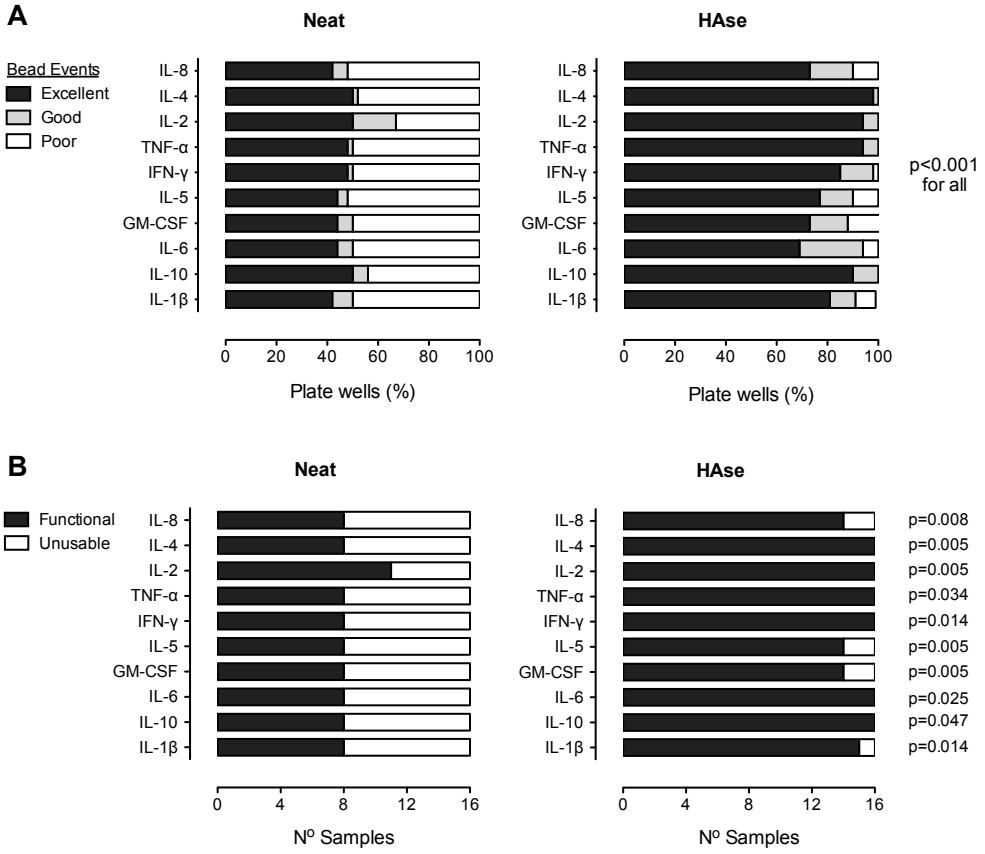


Figure 3.2 – Polystyrene-bead Luminex: Comparison of bead events & functional samples in neat and HAse treated samples. 16 SF samples were run in triplicate for each treatment method. **[A]** The number of wells with poor (<35), good (35-49) and excellent (≥50) bead events for each assay (bead region) when samples were untreated (neat) or HAse treated. The proportion of good or excellent wells was significantly greater after HAse treatment (McNemar's test, $p<0.001$ for each bead region). **[B]** The number of functional (≥2 replicates with ≥35 counts) and unusable samples for each assay. The proportion of functional samples was significantly greater after HAse treatment (McNemar's exact test, $p<0.05$ for each bead region).

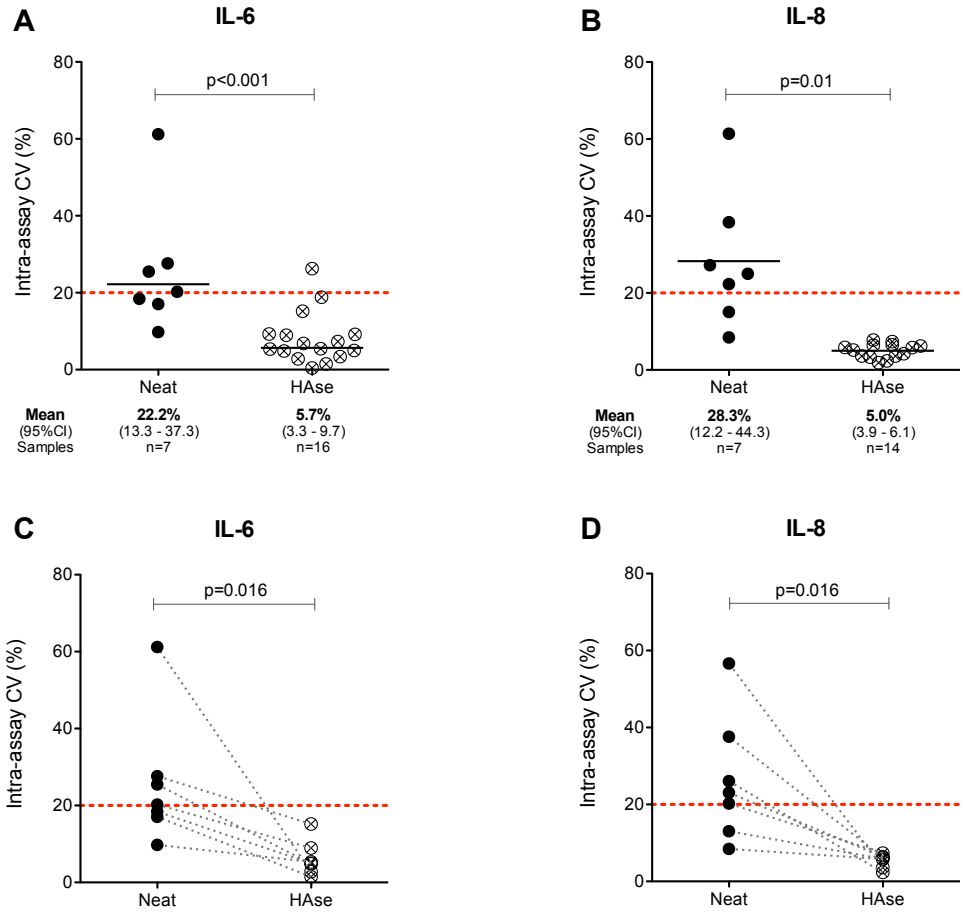


Figure 3.3 – Polystyrene-bead Luminex 10-plex: Comparison of intra-assay precision between neat and HAse treated samples. [A & B] Between-group comparison of intra-assay CV for all functional SF samples was significantly lower in HAse treated than neat samples. **[A]** IL-6 ($p<0.001$, 2-tailed t-test on transformed data, Welch's correction for non-equal variance): X=5.7% (95%CI 3.3-9.7), N=16 vs. XG=22.2% (95%CI 13.3-37.3), N=7; and **[B]** IL-8 ($p=0.01$, 2-tailed t-test with Welch's correction for non-equal variance): X=5.0% (95%CI 3.9-6.1), N=14 vs. X=28.3% (95%CI 12.2-44.3), N=7. X = mean (XG = geometric mean for log transformed data) and 95%CI = 95% confidence interval of the mean. **[C & D]** Within-subject comparison of intra-assay CV in functional samples when neat or HAse treated (N=7) showed a significant reduction in CV after HAse treatment for **[C]** IL-6 ($p=0.016$, paired t-test on transformed data) and **[D]** IL-8 ($p=0.016$, paired t-test). The hashed line at 20% indicates the acceptable threshold for intra-assay CV.

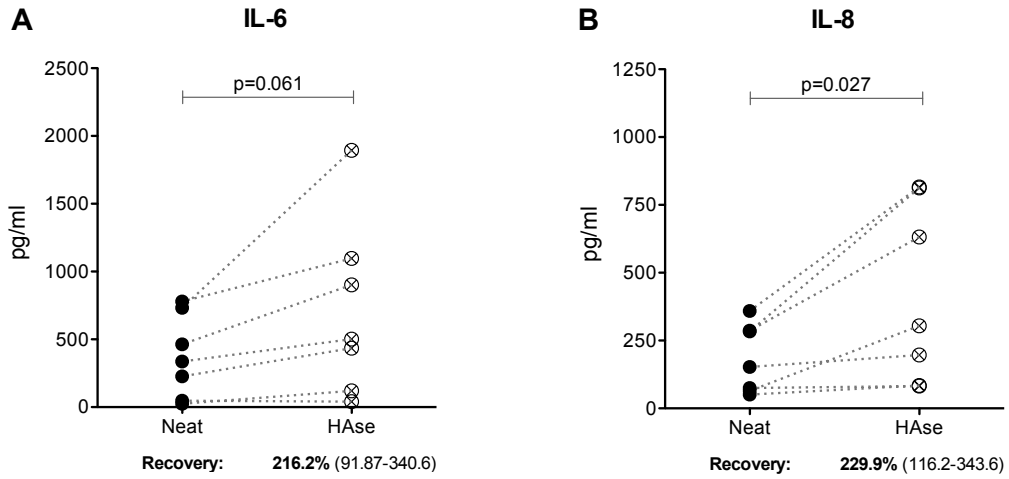


Figure 3.4 – Polystyrene-bead Luminex: Within-subject (pairwise) comparison of dilution-adjusted concentration in functional samples when neat or HAse treated (N=7). HAse treatment resulted in a significantly higher measured concentration of **[B]** IL-8 (p=0.027, paired t-test), but did not reach statistical significance for **[A]** IL-6 (p=0.061, paired t-test on transformed data). The recovery (mean and 95%CI) after HAse treatment is given below each graph.

[Figure 3.5C]. The mean dilution recoveries were elevated, with that for 4-fold dilution (~190%) greater than 2-fold dilution (~150%). However, the 95%CI for both spanned the acceptable range.

3.4.2 MesoScale Discovery Assays

The standard curve points for all 4 analytes met the predefined acceptance criteria. IL-6 and IL-8 were >LOQ in all samples under each treatment. All neat samples achieved an intra-assay CV<20% [Figure 3.6A & B]. Except for one sample in the IL-6 assay, all HAse treated samples also had a CV<20%. Six out of 7 samples had an acceptable CV when 4-fold diluted and 5 out of 7 when 2-fold diluted for both analytes. There was no significant within-subject difference in CV between preparation methods. However, the group average CV for neat and HAse treated samples were lower than diluted samples for both IL-6 and IL-8, but this did not reach statistical significance.

Measured TNF-α levels were in the lower range of the standard curve. Only one sample was above the LOQ after 4-fold dilution: CV of 8.35% and recovery of 62.5%. Concentration and CV data from 4-fold diluted aliquots were therefore not analysed further. Two additional samples fell below quantification: one after 2-fold dilution

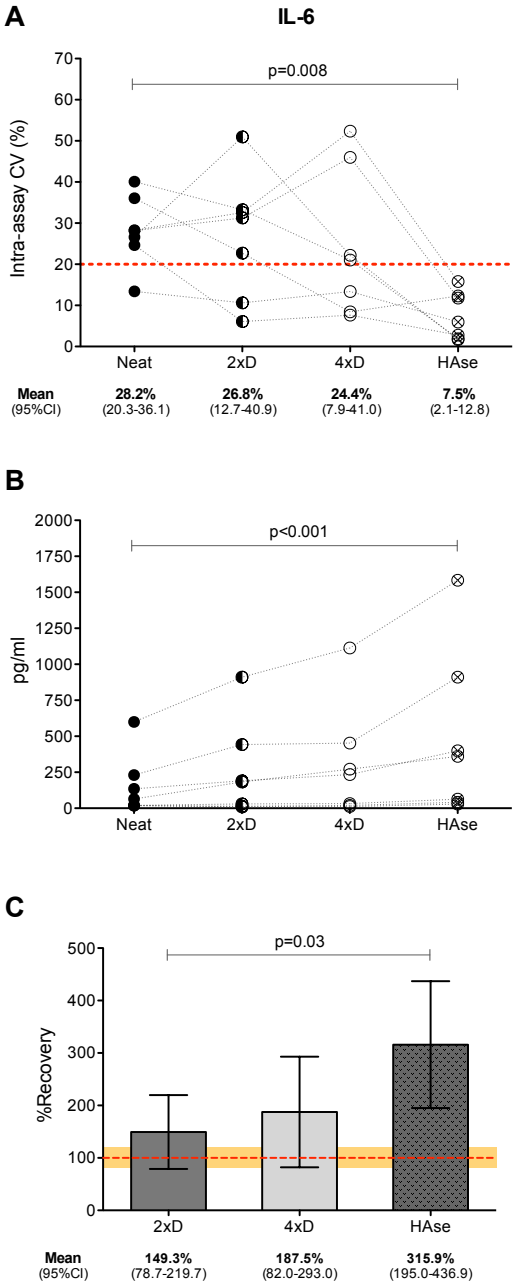


Figure 3.5 – Comparison of [A] intra-assay CV, [B] dilution-adjusted IL-6 concentration and [C] analyte recovery in the magnetic bead Luminex IL-6 assay (N=7). [A] The hashed line at 20% indicates the threshold for acceptable CV. The CV for HAse treated SF was significantly less than when neat or (2- or 4-fold) diluted ($p=0.008$, repeated measures ANOVA; $p<0.05$ for HAse vs. neat, 2-fold and 4-fold dilution by post-hoc Newman-Keuls multiple comparison test). There was no significant difference between neat and diluted samples. Group means and 95%CI are also provided. **[B]** IL-6 concentration was greatest for HAse treated SF ($p<0.001$, repeated measures ANOVA on transformed data; $p<0.05$ for HAse vs. neat, 2-fold and 4-fold dilution by post-hoc Newman-Keuls multiple comparison test). There was no significant difference between neat and diluted samples. **[C]** The hashed line and shading represents the acceptable range for recovery of 100±20%. The IL-6 recovery after HAse treatment was significantly greater than for diluted samples ($p<0.03$, one-way ANOVA; $p<0.05$ for HAse vs. 2-fold and 4-fold dilution by post-hoc Newman-Keuls multiple comparison test). There was no significant difference between diluted samples. Group means and 95%CI are shown.

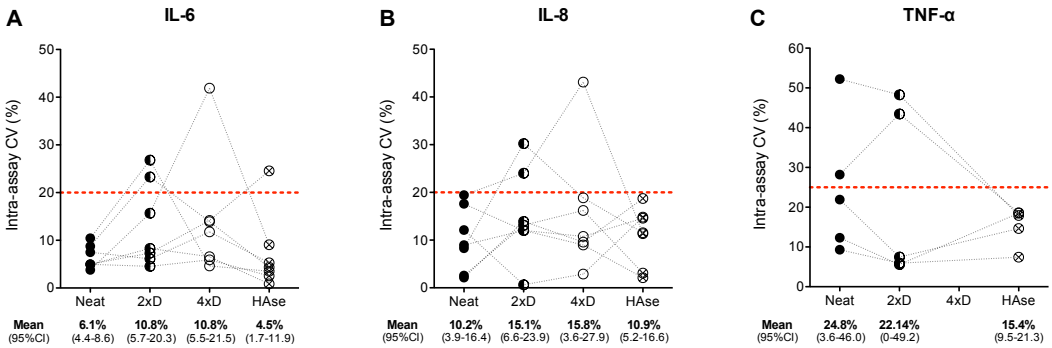


Figure 3.6 – MSD 4-plex: Within-subject (repeat measure) comparison of intra-assay precision. The intra-assay CV for **[A]** IL-6 (p=0.07, repeat measures ANOVA on transformed data, N=7), **[B]** IL-8 (p=0.488, repeat measures ANOVA, N=7), **[C]** TNF-α (p=0.425, repeat measures ANOVA, N=5) showed no significant difference between preparation methods. The hashed line indicates the acceptable threshold for intra-assay CV: 20% for IL-6 and IL-8; 25% for TNF-α (concentrations near LOQ). Group means and 95%CI are provided. Between-group comparisons of CV were not significant.

and another after HAse treatment, leaving 5 samples for within-subject comparisons. An intra-assay CV<25% was achieved in all quantifiable SF samples only after HAse treatment [Figure 3.6C]. However, there was no significant within-subject difference in CV between preparation methods. Using all quantifiable aliquots, the group mean CV after HAse treatment (13.9%, 95%CI 8.0-19.8, n=6) was lower than neat (25.7% 95%CI 12.2-39.2, n=7) and 2-fold (18.5% 95%CI 0-40.9, n=6), but this did not reach statistical significance.

The recoveries for preparation methods are within 20% of expected [Table 3.1] and there were no differences in concentrations of IL-6, IL-8 or TNF-α after adjustment for dilutions [Figure 3.7].

There was only 1 sample with IL-1β levels >LOQ for all 4 preparation methods. A further 3 samples were quantifiable when neat and 2-fold diluted, of which one was also quantifiable when HAse treated. Only one aliquot achieved a CV<25%. The small and irregularly distributed sample size prevented meaningful within-subject or between-group statistical comparisons of endpoints. However, descriptive data on these quantifiable samples is presented in Figure 3.8.

Analyte	Preparation	Recovery (%)	Samples
IL-6	2-fold diluted	106.6 (89.0-124.1)	7 (5)
	4-fold diluted	107.3 (90.4-124.2)	7 (6)
	HAse	103.2 (90.2-116.1)	7 (6)
IL-8	2-fold diluted	105.2 (81.8-128.6)	7 (5)
	4-fold diluted	95.98 (82.6-109.5)	7 (6)
	HAse	87.32 (70.8-103.8)	7 (7)
TNF-α	2-fold diluted	113.7 (73.3-154.1)	6 (4)
	4-fold diluted	-	-
	HAse	83.75 (65.6-102.0)	6 (6)

Table 3.1 – Recoveries of SF treatment methods (w.r.t neat) for MSD Pro-inflammatory 4-Plex. Mean %recovery with 95%CI, and number of samples used to calculate recovery are presented. As a guide to reliability, the number of samples with an acceptable CV is shown in brackets (CV<20% for IL-6 & IL-8, and CV<25% for TNF-α). The acceptable threshold for %recovery is within 100±20% of expected. NB. Only one 4-fold diluted sample was >LOQ for in the TNF-α assay with a recovery of 62.5%.

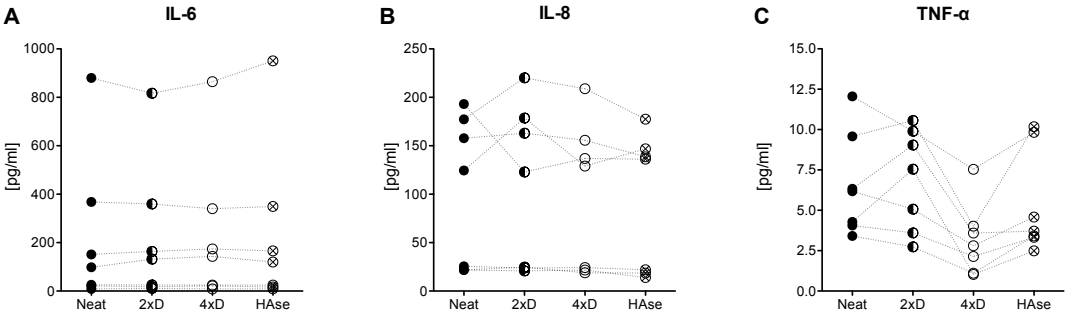


Figure 3.7 – MSD 4-Plex: Repeat measure comparison of dilution-adjusted concentrations. There was no significant difference in adjusted concentrations between preparation methods for **[A]** IL-6 (p=0.668, repeat measures ANOVA on transformed data), **[B]** IL-8 (p=0.509, repeat measures ANOVA), and **[C]** TNF-α (p=0.248, repeat measures ANOVA).

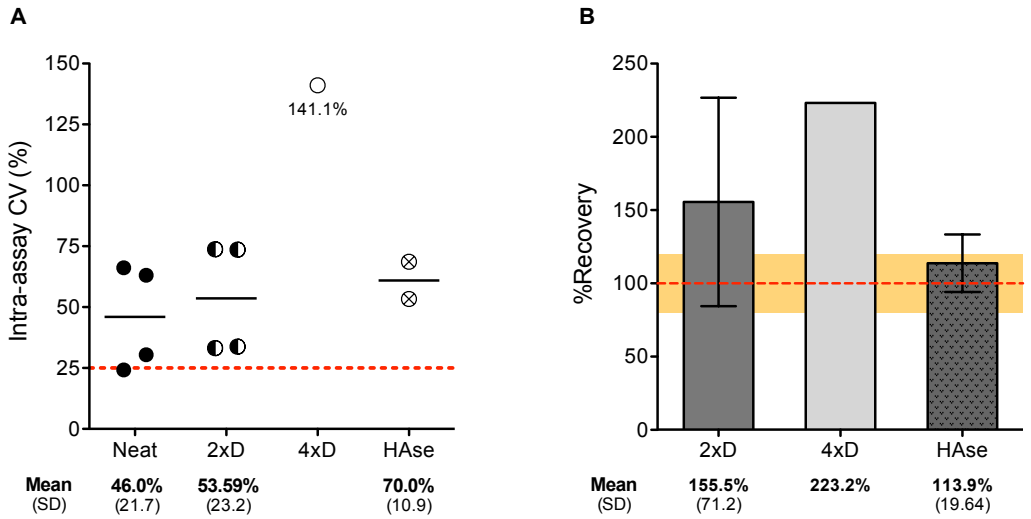


Figure 3.8 – MSD 4-Plex IL-1β assay. **[A]** Intra-assay CV and **[B]** %Recovery of quantifiable samples. Mean and standard deviations (SD) are given. **[A]** The hashed line at 25% indicates the acceptable threshold for intra-assay CV for analytes near LOQ. **[B]** The hashed line and shading represents the acceptable range for recovery of 100±20%.

3.5 Discussion

This study investigated the effect of SF treatment methods on intra-assay precision during analysis by two leading multiplex immunoassay technologies. The experiments demonstrate that HAse treatment can produce more precise concentration measurements than unprepared or diluted SF samples on the Luminex platform, but not conclusively on the MSD platform. In addition, HAse treatment of SF samples improves the measured signal and therefore recovery of target analytes in the bead-based Luminex assays. Finally, HAse treatment improves the efficiency of SF samples to generate useful data in pLuminex assays by increasing the number of plate wells achieving sufficient bead events for analysis.

Testicular HAse breaks down hyaluronan, chondroitin sulphate and keratan sulphate in the SF matrix and is associated with a reduction in protein-matrix interactions [272, 273]. The result is a more homogenous and lower viscosity sample. This reduces pipetting errors thus limiting volume variations that can affect assay accuracy and precision, particularly when working with small volumes. The composition of replicates is more uniform leading to lower variability of measurements.

Pre-treatment of SF is often performed prior to biological analysis. Dilution alone reduces viscosity, but reduces the assay signal and may even render analytes undetectable. None of the cytokines measured in this study were above the dynamic range of their assays even in neat samples. Dilution may place the assay signal in the non-linear region at the low end of the standard curve, thereby increasing the CV [274]. Dithiothreitol (DTT) is an alkylating and reducing agent commonly used in the preparation of other complex biological matrices i.e. sputum [275, 276]. Its properties have also been used to treat SF samples, with a potential additional benefit of reducing interference by heterophilic antibodies [277, 278, 279]. However, DTT can reduce immunoassay detection of several cytokines, chemokines and growth factors since disulphide bond reduction leads to disruption of tertiary structure in the target and possibly the capture antibody [280, 276].

The intra-assay variability in multiplex Luminex assays is also a function of the number of bead events. Higher bead events are associated with greater assay precision, particularly at low ligand concentrations [281, 282]. HAse treatment resulted in higher bead events in the pLuminex 10-plex than neat samples, contributing to the reduction in intra-assay CV. The increase in bead events was specific to SF sample wells; the standard curve wells in both runs achieved maximum bead events per region with acceptable CVs and recoveries. Therefore, the improvement is not an experimental artefact i.e. bead loss during wash steps, and is instead due to a direct effect of HAse on the SF matrix.

Studies on flow-cytometer cell counting in SF samples are of particular relevance to the results on the Luminex platform [283, 270, 271]. These studies demonstrate that HAse treatment improved the precision by reducing errors due to high viscosity related flow. Suguichi et al. illustrated that poor results in untreated samples were associated with cell aggregations [270]. Based on this evidence, one can speculate that HAse treatment reduces bead aggregation and potential bead-matrix interactions. This would create a more uniform distribution of beads in the reaction mixture for the immunoassay to occur. Bead aggregates and bead-matrix complexes could be of variable size and composition (in terms of bead type), and persist after wash steps. They can result in flow errors i.e. clogs during analysis, underestimations of bead events or bead

misclassifications, and retention within the well or filter membrane. The mLuminex IL-6 (monoplex) assay attained excellent bead events irrespective of SF treatment. It is conceivable that the high efficiency of magnetic separation during wash steps prevented the formation and persistence of bead aggregates and complexes, thereby maximising bead events under all treatments. Hence the improved precision in HAse treated samples was due instead to better sample homogeneity and bead dispersal. The greater analyte concentration and considerably elevated recovery in Luminex assays after HAse treatment could be explained by release or exposure of matrix-associated ligands by SF matrix breakdown. Hyaluronic acid is a large polymeric GAG molecule that can physically exclude other molecules (and beads) from a fraction of the solvent volume [208]. With HAse there is an effective increase in matrix "permeability" and "available volume" that offers a larger fluid phase for the ligand capture to occur. In addition, a more uniform bead distribution with elimination of bead aggregates and bead-matrix complexes maximises the surface area for ligand capture and detection antibody binding. The higher assay signals could also contribute to better precision by placing more data points on the linear part of the standard curve. HAse treatment was not assessed alongside dilution in the pLuminex assays. Therefore it is not possible to conclude whether the poor (elevated) recovery is due to the effect of HAse alone, the diluent (from the incorporated 2-fold dilution) or a combination of both. However, for the mLuminex IL-6 assay, the mean recovery after HAse treatment was significantly greater than acceptable and that for 2-fold dilution (>double). Although the mean recoveries for 2-fold and 4-fold dilution in the study were poor (elevated), their 95% confidence intervals spanned the acceptable range of 100±20%. Also in both cases, less than half of samples had an acceptable assay precision. Therefore, one cannot confidently state that there is poor parallelism for Luminex analysis of SF. There is consequently no evidence from this study to suggest that the manufacturer's diluent affects analyte detectability. Taken together, the data supports the notion that HAse activity on the SF matrix is responsible for the improved assay signal on the Luminex platform.

Sandwich immunoassays, and Luminex assays in particular, can be susceptible to the effects of heterophilic antibodies (HAbs) [284, 285]. These are naturally occurring antibodies with low affinities to a variety of antigen targets. Cross-linking of capture

and detection antibodies can result in erroneous results; usually false positive signals. It is possible that SF matrix breakdown by HAse allows greater interference by HAbs contributing to increased measured concentration by Luminex assays. Sample treatments to reduce the potential effects of heterophilic antibodies have been previously described, but can be complicated and costly [286, 287, 288]. Fortunately, the presence of HAbs is likely to be more problematic in the context of inflammatory arthritis rather than OA, and in serum rather than SF. Rheumatoid factor (RF) is often the culprit in inflammatory arthritis. DeForge et al. report RF in 20/23 SF samples from patients with rheumatoid arthritis, but RF was reported as "not available" in the 33 OA SF samples [288]. Likewise, de Jager et al. detected RF in 5/10 rheumatoid arthritis SF samples, but not in any (n=5) OA SF samples [288, 287].

HAse treatment of SF samples during MSD analysis was not as beneficial as for Luminex assays. Neat and HAse treated samples performed similarly with a trend to having better intra-assay precision than diluted samples for analytes measured in the linear region of the standard curve (i.e. IL-6 and IL-8). One can speculate that HAse treatment (with 2-fold dilution) provides better homogeneity of high viscosity SF samples from OA patients than dilution alone, but even so, the average CV for all treatment groups was below 20%. The data suggest that HAse treatment may produce better precision than neat and 2-fold diluted samples for TNF-α, which was detected near the limit of quantification. The improved precision of TNF-α measurements is difficult to explain and may be an artefact of better sample homogeneity with pipetting consistency. The MSD platform, like ELISA, is a solid-phase immunoassay occurring on the base of the specialised plate-wells. It is therefore not subject to the potential fluid-phase and bead-related benefits of HAse treatment postulated earlier for Luminex assays. Sample preparation did not affect the dilution-adjusted concentration and this was reflected in good linearity of dilution (parallelism) and acceptable recovery after treatment. Therefore there is no evidence that the manufacturer's diluent or HAse activity on the SF matrix affects analyte detectability. The data are in alignment with the manufacturer's claims that MSD assays have a high sensitivity, wide dynamic ranges and are robust to matrix effects. However, SF matrix interference and response to HAse would need to be formally assessed by spike-and-recovery experiments in both MSD and Luminex platforms.

At the time of writing, there were no previous investigations comparing SF preparation methods for multiplex immunoassay analysis. Although the number of SF samples in this study was small, every sample was tested under each preparation method. The within-subject experimental design, with matched (repeated measures) comparisons, improves the study power. However, there are a number of limitations to this study. The intra-assay CV was calculated from only triplicate measurements; 5 or more technical replicate would be more reliable. Triplicates were used to preserve precious small volume OA samples and maximise the number of samples tested on an each assay plate and thus cost-efficiency. Ideally the standard curve should be prepared in the same matrix as samples tested. It was not feasible to prepare standard curves in the appropriate matrix for every preparation method in each comparison experiment. There was no single large volume OA sample for preparation of multiple standard curves and I wanted to avoid pooling of precious samples. Therefore, a pragmatic approach was taken by making dilutions and HAse preparation in the same diluent used for the standard curve preparation [263]. An arbitrary final concentration of 2mg/ml HAse equivalent to ~1200 U/ml was used. Further work to compare different HAse concentrations would be valuable. Similar improvements in Luminex assay performance achieved with lower concentrations would allow more efficient use of consumables. Equally, the prospect that MSD assays may require higher concentrations for performance benefits needs to be tested. The incorporated dilution involved with HAse treatment (2-fold in this study) warrants further assessment, particularly for Luminex assays: lower built-in dilutions could promote assay signals falling even higher on the standard curve, and HAse may prevent samples falling below the detection limit at higher dilutions. The study did not test inter-assay precision, which is influenced by numerous factors: reagent stability, errors in reagent preparation; variations in pH of buffers, temperature, wash procedure; pipetting errors and instrument calibration/performance fluctuations. Sufficient numbers of SF samples with enough volume for numerous runs under each treatment protocol were not available to examine these factors in a within-subject experimental design.

3.6 Conclusions

Both null hypotheses are rejected for the Luminex platform, but neither can be rejected for the MSD platform.

The study was motivated by poor and variable cytokine measurements in neat SF samples from patients with OA. The data demonstrated that HAse treatment of SF samples permits better precision and higher assay signals during multiplex analysis on the Luminex platform compared to neat or diluted samples. These performance advantages may be a result of increased efficiency of bead-based sandwich immunoassays in the fluid phase. One can, therefore, recommend routine HAse treatment of SF samples prior to analysis on the Luminex platform. HAse treatment conferred no conclusive benefits to assay precision or signal on the solid-phase MSD platform, but neither were there any detrimental consequences. Characterisation of low volume SF samples from OA patients will require analysis on multiple platforms to gain the maximum biological information. These analyses may often have to be performed simultaneously to avoid sample freeze-thaw cycling, especially if adequate multiple aliquots were not possible after sample collection. For these reasons, it is possible to advocate routine HAse treatment to reduce SF sample viscosity for biomarker research in OA.

This chapter has been published as:

Jayadev C, Rout R, Price A, Hulley P, Mahoney D. Hyaluronidase treatment of synovial fluid to improve assay precision for biomarker research using multiplex immunoassay platforms. J Immunol Methods. 2012 Dec 14;386(1-2):22–30.

Chapter 4

Multiplex Analysis of Synovial Fluid: A Comparison of Immunoassay Platforms

4.1 Introduction

Multiplex immunoassay platforms are indispensable for biomarker discovery. They allow high-throughput, multiple parallel analysis of small volume biological specimens that are difficult to obtain. The two dominant technologies are Luminex and MesoScale Discovery (MSD), both of which are sandwich immunoassays. The fundamental difference is that Luminex assays occur in the fluid-phase on microspheres (beads), whilst the MSD assays occur in the solid-phase on a microdot array at the floor of the plate-well. Neither platform is validated for synovial fluid (SF) analysis according to their manufacturer's protocols. There are very few studies in the literature reporting their use with SF, particularly in the context of osteoarthritis (OA).

"The utility of biomarkers to support clinical decision making in clinical studies, or to conduct exploratory studies, is directly related to the quality of the underlying bioanalytical data." [263]. In Chapter 3, hyaluronidase treatment of SF was shown to improve intra-assay precision and signal measurement in Luminex assays, but there was no conclusive benefit for MSD assays that were relatively resilient to matrix effects. Additional points were raised but not explored. For example, TNF-α was quantifiable in the SF samples by the MSD assay but not the Luminex assay. The magnetic-bead

Luminex assay was a monoplex and its performance in terms of bead events/region and intra-assay precision may not generalise to a multiplex assay.

Cross-comparison of multiplex formats and comparison against the existing "gold standard" of ELISA are required to facilitate quality control and guide optimum assay selection for biomarker discovery. Results of previous comparison and validation studies in culture medium, serum and plasma cannot be extrapolated to the complex biological matrix of synovial fluid [289, 290, 291, 292, 293, 294].

This chapter makes a comprehensive and detailed comparison between Luminex and MSD platforms during the analysis of SF samples from patients with end-stage knee OA. Both versions of the Luminex format i.e. polystyrene and magnetic beads are assessed. Assay performance parameters that are important for data quality control and the agreement of concentration measurements between platforms are examined.

4.2 Null Hypotheses

Analysis of SF from patients with end-stage knee OA by the Luminex and MSD platforms demonstrate:

1. Acceptable assay performance in terms of limits of detection and quantification, intra-assay precision and spike recovery
2. Equivalent assay performance between platforms in terms of limits of detection and quantification, intra-assay precision and spike recovery
3. Absolute agreement of concentration measurements between platforms assessed by intraclass correlation, concordance, weighted Deming regression and Bland-Altman analysis.

4.3 Methodology

SF was collected from 31 patients with end-stage knee OA at the time of primary joint replacement surgery. Samples were harvested, processed and stored as previously outlined. Prior to analysis, thawed and clarified aliquots were treated with 2mg/ml bovine testicular hyaluronidase (type I-S, 618.4 U/mg, Sigma) as detailed in Chapter 3.

4.3.1 Immunoassays

SF aliquots from each patient (N=31) were analysed by a polystyrene-bead Luminex (pLuminex) human cytokine 25-plex, a magnetic-bead Luminex (mLuminex) human cytokine 25-plex and a MSD pro-inflammatory 4-plex. Both Luminex assays were from the same manufacturer (Life Technologies). On all platforms, calibrators and blanks were measured in duplicate and SF aliquots were run in triplicate.

ELISA for IL-6 and TNF-α was performed on SF aliquots from patients previously analysed by the pLuminex 25-plex (n=10), mLuminex 25-plex (n=8) and MSD 4-plex (n=11). IL-6 and TNF-α were chosen based on observations in Chapter 3, where the former was readily quantifiable in all platforms and the latter was only quantifiable on the MSD platform.

Measured sample concentrations above the limit of quantification (LOQ) were adjusted for the 2-fold dilution inherent to hyaluronidase treatment of SF. The experimental plan is summarised in Figure 4.1.

4.3.2 Sample Spiking

Two aliquots of one SF sample to be run on the mLuminex and MSD platforms were spiked with known concentrations of calibrator from the standard curve of the respective assays. The expected concentrations for MSD spiked aliquots were 833.3pg/ml (spike-A) and 208.3pg/ml (spike-B) for all analytes. The concentrations of each analyte in the Luminex stock calibrator are different and spiked sample concentrations are given in Table 4.1. All spiked samples were measured in triplicate.

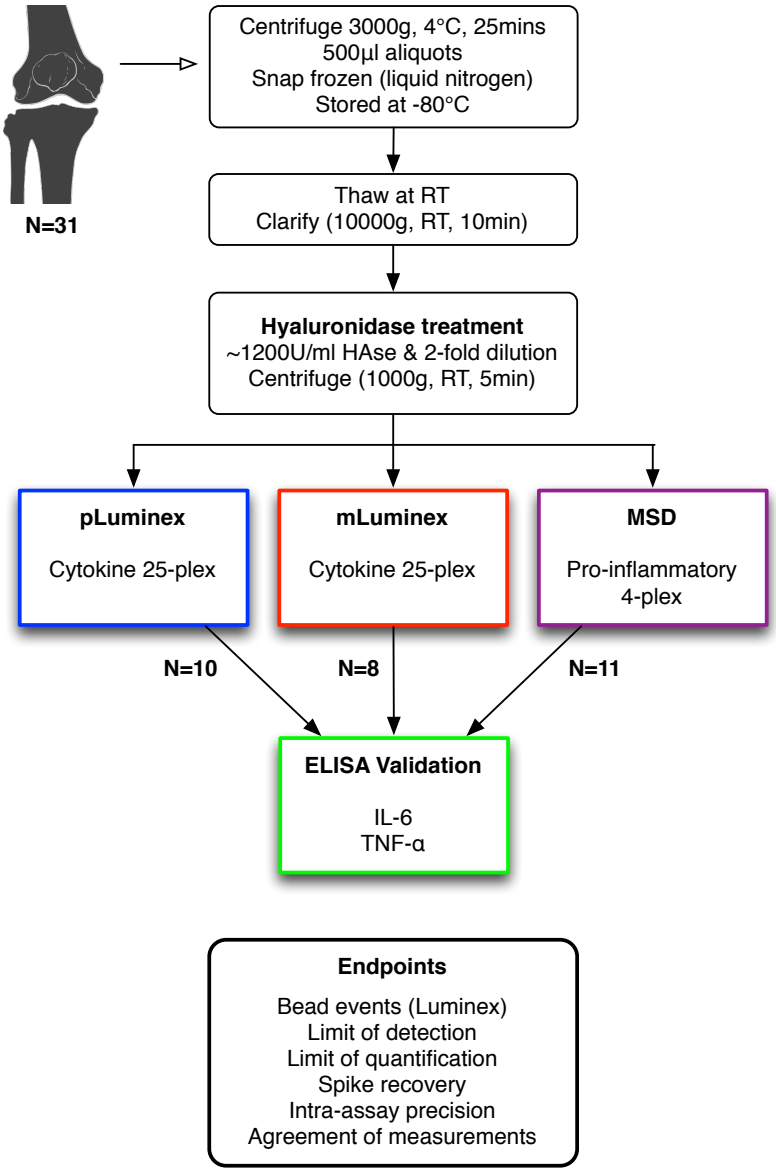


Figure 4.1 – Within-subject experimental design for comparison and validation of Luminex and MSD platforms. 31 SF samples were harvested and underwent identical processing. Each sample was analysed by a polystyrene-bead Luminex (pLuminex) cytokine 25-plex (IL1-β, IL-6, TNF-α, IL-4, IL-8, IL-10, IL-13, IL-12, RANTES, Eotaxin, IL-17, MIP-1α, GM-CSF, MIP-1β, MCP-1, IL-15, IL-5, IFN-γ, IFN-α, IL-2, IL-7, IP-10, IL-2R, MIG, IL-4); magnetic-bead Luminex (mLuminex) cytokine 25-plex (same analytes as pLuminex), and a MesoScale Discovery (MSD) pro-inflammatory 4-plex (IL1-β, IL-6, IL-8 and TNF-α). Each platform was validated against ELISA for IL-6 & TNF-α. The platforms were compared for bead events (Luminex only), limits of detection and quantification, spike recovery, and intra-assay replicate CV. The agreement of cytokine measurements by each platform was tested.

Assay	Spike-A	Spike-B
IL-1β	691.7	76.9
IL-6	345.8	38.4
IL-8	1066.7	118.5
TNF-α	725.0	80.6
IL-10	1825.0	202.8
IL-13	1691.7	188.0
IL-12	541.7	60.2
RANTES	291.7	32.4
Eotaxin	241.7	26.9
IL-17	3233.3	359.3
MIP-1α	2083.3	231.5
GM-CSF	1445.8	160.6
MIP-1β	600.0	66.7
MCP-1	508.3	56.5
IL-15	2025.0	225.0
IL-5	194.2	21.6
IFN-γ	875.0	97.2
IFN-α	1075.0	119.4
IL-1RA	2333.3	259.3
IL-2	100.0	11.1
IL-7	816.7	90.7
IP-10	75.0	8.3
IL-2R	1750.0	194.4
MIG	200.0	22.2
IL-4	3550.0	394.4

Table 4.1 – Spiked concentrations (pg/ml) for the mLuminex human cytokine 25-plex assay.

4.3.3 Assay Performance Parameters

Luminex Bead Events

The data from SF samples were defined as "ideal" if at least 2 replicates had $\geq$50 bead events/region and "functional" if at least 2 replicates had $\geq$35 bead events/region; otherwise they were defined as "unusable". Only measured concentrations from "functional" or "ideal" samples were used for further analyses.

The proportion of functional and ideal samples for each analyte was compare for the mLuminex and pLuminex platforms using the McNemar (exact) test. The Wilcoxon matched-pairs signed rank test was used to compare the average (median) number of functional and ideal samples between bead systems.

Limit of Detection

Analyte LODs were derived using duplicate blank measurements from 3 pLuminex, 8 mLuminex and 5 MSD plate runs [Appendix C]. The LOD for an analyte was described by the geometric mean and 95% confidence interval (95%CI) from the multiple plate runs. Platforms were compared using the 2-tailed t-test or one-way ANOVA. For each analyte, the proportion of samples above LOD on each platform was compared using either McNemar's test (2 matched proportions) or Cochran's Q test (3 matched proportions).

Limit of Quantification

Analyte LOQs were derived from the standard curves of 3 pLuminex, 8 mLuminex and 5 MSD plate runs [Appendix C]. The LOQ for an analyte was described by the median and interquartile range (IQR) from the serial plate runs. The median was used because the calibrator points of the standard curve are arbitrary discrete concentrations. Platforms were compared using the Mann-Whitney U-test or Kruskal-Wallis test. For each analyte, the proportion of samples above LOQ on each platform was compared

using either McNemar's test (2 matched proportions) or Cochran's Q test (3 matched proportions).

Spike Recovery

The percentage recovery in spiked samples was calculated using the formula given in Appendix A.2. The expected known spike concentrations were as given above. Measured concentrations falling below the LOD or LOQ were substituted with the $^{LOD}/_{2}$ and $^{LOQ}/_{2}$ respectively. Recoveries within ±20% of expected were considered acceptable.

Intra-Assay Precision

The coefficient of variation (CV) for replicate measurements was used as a measure of intra-assay precision. The CV was compared between platforms using the Wilcoxon signed-rank test or Friedman test (with post-hoc Dunn's multiple comparison test). Non-parametric methods were used since the presence of zero values prevented routine transformations (i.e. logarithms) without data loss. The number of samples with a replicate CV $\leq$20% was compared per analyte using either McNemar's test or Cochran's Q test. Platform comparisons were only made for analytes with at least 10 (out of 31) matched sample measurements >LOD.

4.3.4 Cross-Platform Agreement

The agreement of analyte concentration measurements for the 31 SF samples between platforms was assessed by 4 methods:

- Intraclass correlation coefficient (ICC) [295]
- Concordance correlation coefficient (CCC) [296, 297]
- Weighted Deming regression [298, 299, 300]
- Bland-Altman analysis [301, 302]

SF measurements for IL-1β, TNF-α, IL-6 and IL-8 were made by all three platforms. The CCC, Bland-Altman analysis and Deming regression assess agreement between 2 methods and were therefore calculated for each platform pairing: pLuminex vs. mLuminex, pLuminex vs. MSD and mLuminex vs. MSD. The ICC was calculated for all 3 platforms together and for individual pairings. Skewed data were $\log_{10}$-transformed and confirmed for a normal distribution prior to calculations of ICC and CCC.

The ICC was used to measure the absolute agreement and consistency (relative agreement) between platforms using a two-way random effects model for individual sample measurements, $ICC_{2,1}$[303]. The ICC and 95% confidence interval are presented. There are many arbitrary scales for interpretation of the ICC, of which the following was used in this study:

Value of ICC	Strength of Agreement
<0.2	Poor
0.2 - 0.5	Fair
0.5 - 0.7	Moderate
0.7 - 0/8	String
>0.8	Almost Perfect

Negative ICCs are not theoretically possible, but statistical software can output them. Typically this happens when the between-subject variation is relatively small compared to the within-subject variation (due to different raters). Negative ICC estimates suggest poor agreement, but cannot be reliably interpreted.

The CCC was interpreted using the scale [304]:

Value of CCC	Strength of Agreement
<0.90	Poor
0.90 - 0.95	Moderate
0.95 - 0.99	Substantial
>0.99	Almost perfect

Weighted Deming regression was performed since the standard deviation of replicate measurements was not constant. Instead a constant coefficient of variation (CV) is assumed and the statistical software was provided with the pooled CV for each platform calculated from the CV of each sample's triplicate measurements. Jack-knife resampling

was used to estimate the standard errors of the slope and intercept. Therefore, the requirement that the analytical differences are normally distributed is not required [298]. The slope (proportional error) and intercepts (systematic error) of the regression model are presented with their respective 95% confidence intervals.

Analytes with fewer than 15 (out of 31) matched measurements below LOQ were excluded from the repeatability analysis.

4.4 Results

4.4.1 Luminex Bead Events

The mLuminex 25-plex achieved $\geq$100 bead events/region in all sample replicate wells. Consequently, all 31 samples were categorised as ideal (and thus functional) for each analyte (bead region). There were no analytes on the pLuminex 25-plex for which all samples were ideal samples and only 3 analytes (RANTES, IL1-RA and IL-4) for which all samples were functional [Table 4.2 & Table 4.3].

The number of ideal samples in the mLuminex 25-plex was significantly greater than the pLuminex 25-plex for 15 analytes. This remained significant for 6 analytes after adjustment for multiple comparisons. There were significantly more functional samples for 5 analytes in the mLuminex 25-plex, but only one after multiple comparison correction. The overall median number of ideal and functional samples in the pLuminex 25-plex was 24 (77%, IQR 70 – 87%) and 28 (90%, IQR 84 – 97%) respectively. These were both significantly less than the mLuminex 25-plex ($p<0.001$).

4.4.2 Limits of Detection

The MSD platform had the lowest LOD for the 4 analytes in common with the Luminex assays [Table 4.4]. This was significantly lower than mLuminex for all 4 analytes and lower than pLuminex for IL-1β and IL-6 [Figure 4.2A]. The LODs for the pLuminex were lower than mLuminex for all analytes except IP-10. This was statistically significant

	mLuminex		pLuminex			
Assay	**N**	**%**	**N**	**%**	**p-value**	
IL-4	31	100%	30	97%	1.000	
RANTES	31	100%	29	94%	0.500	
IL-1RA	31	100%	29	94%	0.500	
IP-10	31	100%	29	94%	0.500	
IL-10	31	100%	28	90%	0.250	
MIP-1β	31	100%	27	87%	0.125	
TNF-α	31	100%	27	87%	0.125	
IL-2R	31	100%	27	87%	0.125	
MIP-1α	31	100%	26	84%	0.063	
IL-2	31	100%	26	84%	0.063	
IL-15	31	100%	24	77%	0.016	*
GM-CSF	31	100%	24	77%	0.016	*
IFN-γ	31	100%	24	77%	0.016	*
IL-6	31	100%	23	74%	0.008	*
IL-1β	31	100%	23	74%	0.008	*
IL-12	31	100%	23	74%	0.008	*
IL-5	31	100%	23	74%	0.008	*
IL-8	31	100%	22	71%	0.004	*
IL-17	31	100%	22	71%	0.004	*
IFN-α	31	100%	21	68%	0.002	*
IL-13	31	100%	21	68%	0.002	**
MIG	31	100%	20	65%	0.001	**
Eotaxin	31	100%	18	58%	<0.001	**
IL-7	31	100%	16	52%	<0.001	**
MCP-1	31	100%	13	42%	<0.001	**
Median	**31**	**100%**	**24**	**77%**	**<0.001**	

Table 4.2 – Comparison of the number of ideal samples for each assay (bead region) on the mLuminex and pLuminex 25-plex. Ideal samples: ≥50 bead events/region for ≥2 replicates. All 31 samples were ideal in the mLuminex 25-plex. P-values are given for paired comparison of proportions (McNemar's exact test). Asterisks (*) denote significance (p<0.05) and a double asterisk (**) denotes significance (q<0.05) after correction for multiple comparisons by Holm's method. The median number of ideal samples was compared using the Wilcoxon matched-pairs signed rank test.

	mLuminex		pLuminex			
Assay	**N**	**%**	**N**	**%**	**p-value**	
IL-4	31	100%	31	100%	1.000	
RANTES	31	100%	31	100%	1.000	
IL-1RA	31	100%	31	100%	1.000	
IP-10	31	100%	30	97%	1.000	
IL-10	31	100%	30	97%	1.000	
MIP-1β	31	100%	30	97%	1.000	
TNF-α	31	100%	30	97%	1.000	
IL-2R	31	100%	29	94%	0.500	
MIP-1α	31	100%	30	97%	1.000	
IL-2	31	100%	30	97%	1.000	
IL-15	31	100%	28	90%	0.250	
GM-CSF	31	100%	28	90%	0.250	
IFN-γ	31	100%	28	90%	0.250	
IL-6	31	100%	30	97%	1.000	
IL-1β	31	100%	27	87%	0.125	
IL-12	31	100%	27	87%	0.125	
IL-5	31	100%	27	87%	0.125	
IL-8	31	100%	26	84%	0.063	
IL-17	31	100%	24	77%	0.016	*
IFN-α	31	100%	27	87%	0.125	
IL-13	31	100%	25	81%	0.031	*
MIG	31	100%	26	84%	0.063	
Eotaxin	31	100%	23	74%	0.008	*
IL-7	31	100%	22	71%	0.004	*
MCP-1	31	100%	19	61%	0.001	**
Median	**31**	**100%**	**28**	**90%**	**<0.001**	

Table 4.3 – Comparison of the number of functional samples for each assay (bead region) on the mLuminex and pLuminex 25-plex. Functional samples: ≥35 bead events/region for ≥2 replicates. All 31 samples were ideal (and therefore functional) in the mLuminex 25-plex. P-values are given for paired comparison of proportions (McNemar's exact test). Asterisks (*) denote significance ($p<0.05$) and a double asterisk (**) denotes significance ($q<0.05$) after correction for multiple comparisons by Holm's method. The median number of functional samples was compared using the Wilcoxon matched-pairs signed rank test.

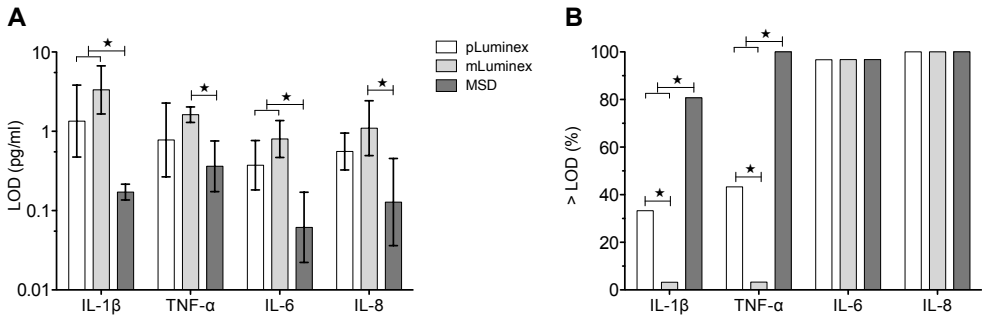


Figure 4.2 – Comparison of [A] LOD and [B] Number of samples >LOD for IL-1β, TNF-α, IL-6 and IL-8 between pLuminex, mLuminex and MSD platforms. The LOD concentration is on a log scale and the geometric mean and 95%CI are shown in [A]. The asterisk denotes statistical significance ($p<0.05$).

for 6 analytes, but only 3 after correction for multiple comparisons (IL-10, MIP-1α and IL-4) [Table 4.4].

Six analytes (IL-8, IL-12, Eotaxin, MCP-1, IL-15 and IP-10) were detectable in all samples irrespective of platform. IL-6 was >LOD in all but one (same) sample on all 3 platforms. Two analytes (IL-5 and IFN-γ) were undetectable in the 31 samples on both Luminex platforms. For 10 of the remaining analytes (6 after multiple comparison adjustment), the proportion of samples >LOD was significantly different between platforms [Table 4.5 and Figure 4.2B]. This largely mirrored the platform with the lowest LOD concentrations. In particular, the number of samples for which IL-1β and TNF-α were detectable was much greater on the MSD platform (>80%).

There were 4 anomalous results. GM-CSF, IFN-α and IL-2R had (non-significantly) lower LODs on the pLuminex, but were detected in significantly more samples by the mLuminex. MIP-1α had a significantly lower LOD on the pLuminex, but was detected in (non-significantly) more samples by the mLuminex.

4.4.3 Limit of Quantification

All 4 MSD assays had a LOQ of 0.61pg/ml. This was lower than the corresponding Luminex assays, but did not reach statistical significance for IL-6 [Table 4.6 and Figure 4.3A]. The LOQ was equivalent on both Luminex platforms for 12 analytes. For 11

	pLuminex		mLuminex		MSD			
Assay	**LOD**	**95%CI**	**LOD**	**95%CI**	**LOD**	**95%CI**	**p-value**	
IL-1β	1.35	0.47-3.28	3.33	1.66-6.68	**0.17**	0.14-0.22	<0.001	**
IL-6	0.37	0.18-0.76	0.80	0.47-1.37	**0.06**	0.02-0.17	<0.001	**
IL-8	0.56	0.33-0.95	1.10	0.50-2.43	**0.13**	0.04-0.45	0.004	**
TNF-α	0.78	0.27-2.27	1.62	1.29-2.03	**0.36**	0.17-0.76	<0.001	**
IL-10	**0.12**	0.06-0.23	1.62	1.22-2.16			<0.001	**
IL-13	**1.35**	0.04-40.94	6.04	3.20-11.40			0.174	
IL-12	**1.51**	0.33-6.96	3.36	1.94-5.83			0.112	
RANTES	**3.59**	1.15-11.28	7.35	4.60-11.70			0.071	
Eotaxin	**0.06**	0.01-0.45	0.21	0.12-0.35			0.079	
IL-17	**2.52**	0.25-25.63	11.91	8.26-17.20			0.079	
MIP-1α	**4.06**	2.14-7.69	10.51	8.26-13.20			0.002	**
GM-CSF	**2.65**	0.76-9.23	6.26	4.40-8.91			0.055	
MIP-1β	**4.39**	0.90-21.24	13.76	9.09-20.80			0.050	
MCP-1	**1.66**	0.63-4.37	2.45	1.80-3.32			0.199	
IL-15	**1.09**	0.30-4.00	6.42	4.09-10.10			0.004	*
IL-5	**0.44**	0.05-4.10	0.95	0.72-1.26			0.264	
IFN-γ	**0.43**	0.03-6.85	1.69	1.19-2.41			0.150	
IFN-α	**7.79**	1.34-45.12	9.05	5.15-15.90			0.764	
IL-1RA	**9.00**	4.73-17.15	23.78	11.80-47.80			0.015	*
IL-2	**0.45**	0.06-3.64	0.95	0.54-1.65			0.245	
IL-7	**3.94**	0.89-17.40	16.71	11.90-23.50			0.024	*
IP-10	0.83	0.38-1.84	**0.56**	0.38-0.81			0.141	
IL-2R	**7.59**	4.39-13.13	11.73	6.95-19.80			0.117	
MIG	**4.60**	0.80-26.57	7.95	4.03-15.70			0.315	
IL-4	**1.40**	0.74-2.63	5.09	4.25-6.10			0.001	**

Table 4.4 – Limits of detection (LOD) for each analyte and platform. The geometric mean and 95%CI in pg/ml are shown. LOD calculations were based on blank signals from a series of pluminex (3 plates), mLuminex (8 plates), and MSD (5 plates) experiments. The lowest LOD for each analyte is given in bold. Log transformed LODs for each analyte were compared across platforms, using one-way ANOVA for IL-1β, IL-6, IL-8 & TNF-α, and two-tailed t-test for the remaining analytes. Asterisks (*) denote significance ($p<0.05$) and a double asterisk (**) denotes significance ($q<0.05$) after correction for multiple comparisons by Holm's method.

	pLuminex		mLuminex		MSD			
Assay	**%**	**n/N**	**%**	**n/N**	**%**	**n/N**	**p-value**	
IL-1β	33.3	9/27	3.2	1/31	**80.7**	25/31	<0.001	**
IL-6	96.7	29/30	96.8	30/31	96.8	30/31		
IL-8	100.0	26/26	100.0	31/31	100.0	31/31		
TNF-α	43.3	13/30	3.3	1/31	**100.0**	31/31	<0.001	**
IL-10	**33.3**	10/30	9.7	3/31			0.039	*
IL-13	**40.0**	10/25	6.5	2/31			0.022	*
IL-12	100.0	27/27	100.0	31/31				
RANTES	**83.9**	26/31	45.2	14/31			0.001	**
Eotaxin	100.0	23/23	100.0	31/31				
IL-17	12.5	3/24	**16.1**	5/31			1.000	
MIP-1α	40.0	12/30	**51.6**	16/31			0.453	
GM-CSF	0	0/28	**25.8**	8/31			0.008	*
MIP-1β	**93.3**	28/30	38.7	12/31			<0.001	**
MCP-1	100.0	19/19	100.0	31/31				
IL-15	100.0	28/28	100.0	31/31				
IL-5	0.0	0/27	0.0	0/31				
IFN-γ	0.0	0/28	0.0	0/31				
IFN-α	14.8	4/27	**54.8**	17/31			0.001	**
IL-1RA	**71.0**	22/31	61.3	19/31			0.453	
IL-2	**20.7**	6/29	9.7	3/31			0.508	
IL-7	**86.4**	19/22	9.7	3/31			<0.001	**
IP-10	100.0	30/30	100.0	31/31				
IL-2R	62.1	18/29	**83.9**	26/31			0.039	*
MIG	**30.8**	8/26	22.6	7/31			1.000	
IL-4	0.0	0/31	0.0	0/31				

Table 4.5 – Proportions of SF samples (N=31) above the limit of detection (LOD) for each analyte and platform given in Table 4.4. The highest proportion for each analyte is given in bold. Matched proportions were compared by McNemar's exact test or Cochran's Q test. Asterisks (*) denote significance (p<0.05) and a double asterisk (**) denotes significance (q<0.05) after correction for multiple comparisons by Holm's method.

	pLuminex		mLuminex		MSD			
Assay	**LOQ**	**IQR**	**LOQ**	**IQR**	**LOQ**	**IQR**	**p-value**	
IL-1β	3.80	3.80-11.39	7.60	3.80-11.39	**0.61**	0.61-2.44	0.005	**
IL-6	1.90	1.90-1.90	1.90	1.90-1.90	**0.61**	0.38-2.44	0.751	
IL-8	5.85	5.85-5.85	5.85	5.85-5.85	**0.61**	0.38-1.53	0.002	**
TNF-α	3.98	3.98-3.98	3.98	3.98-3.98	**0.61**	0.61-2.44	0.002	**
IL-10	10.01	10.01-10.01	10.01	10.01-10.01			1.000	
IL-13	9.28	9.28-9.28	9.28	9.28-23.21			0.241	
IL-12	**2.97**	2.97-8.92	5.95	2.97-22.29			0.500	
RANTES	**4.8**	4.8-14.40	14.40	14.40-14.40			0.027	*
Eotaxin	1.33	1.33-1.33	1.33	1.33-1.33			0.364	
IL-17	**17.74**	17.74-17.74	35.48	17.74-53.28			0.155	
MIP-1α	11.43	11.43-11.43	11.43	11.43-34.29			0.236	
GM-CSF	**7.93**	7.93-23.80	23.80	11.90-23.80			0.223	
MIP-1β	**9.88**	9.88-9.88	29.63	29.63-29.63			0.010	*
MCP-1	8.37	2.79-9.88	**5.58**	2.79-8.37			0.367	
IL-15	11.11	11.11-11.11	11.11	11.11-27.78			0.361	
IL-5	3.20	1.07-11.11	**2.14**	1.07-3.20			0.367	
IFN-γ	4.80	4.80-4.80	4.80	4.80-4.80			1.000	
IFN-α	**11.93**	5.09-53.09	17.70	8.85-44.24			0.526	
IL-1RA	38.41	12.80-38.41	38.41	38.26-115.20			0.329	
IL-2	4.94	4.94-4.94	**3.30**	1.65-4.94			0.151	
IL-7	**13.44**	13.44-40.33	40.33	40.33-40.33			0.087	
IP-10	1.23	1.23-1.23	1.23	1.23-1.23			0.697	
IL-2R	**9.60**	9.60-28.81	28.81	28.81-28.81			0.087	
MIG	**9.88**	3.29-29.63	29.63	12.57-29.63			0.314	
IL-4	19.48	19.48-19.48	19.48	19.48-19.48			1.000	

Table 4.6 – Limits of quantification (LOQ) for each analyte and platform. The median and IQR in pg/ml are shown. LOQs were derived from the standard curves of a series of pluminex (3 plates), mLuminex (8 plates), and MSD (5 plates) experiments. The lowest LOQ for each analyte is given in bold. LOQs for each analyte were compared across platforms, using the Kruskal-Wallis test for IL-1β, IL-6, IL-8 and TNF-α, and Mann-Whitney U-test for the remaining analytes. Asterisks (*) denote significance ($p<0.05$) and a double asterisk (**) denotes significance ($q<0.05$) after correction for multiple comparisons by Holm's method.

analytes the LOQ was lower on the polystyrene than magnetic Luminex system. This was significant for only RANTES and MIP-1, but significance was lost if corrected for multiple-comparisons [Table 4.6]. The LOQ was lower on the mLuminex for 3 analytes, but this was not statistically different.

IL-6 and IL-8 measurements were >LOQ in most (>88%) of the osteoarthritic SF samples tested on all three platforms [Table 4.7]. IL-1β and TNF-α levels were <LOQ on nearly all samples on both Luminex platforms and were quantifiable in significantly more samples on the MSD platform [Figure 4.3B]. However, IL-1β was >LOQ in less than 1/3 of samples tested even on the MSD platform.

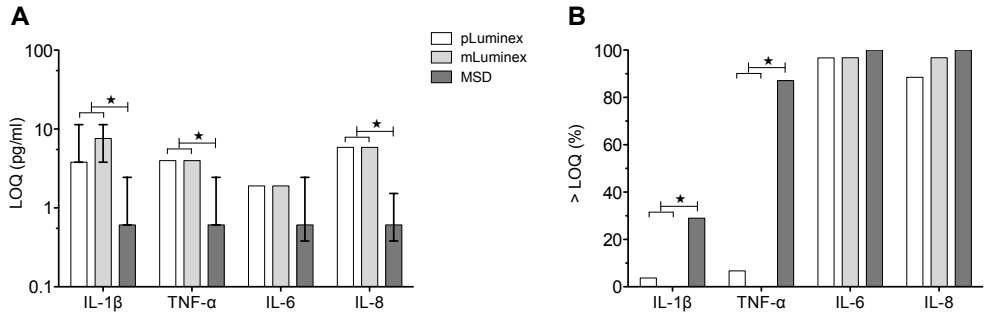


Figure 4.3 – Comparison of [A] LOQ and [B] Number of samples >LOQ for IL-1β, TNF-α, IL-6 and IL-8 between pLuminex, mLuminex and MSD platforms. The LOQ concentration is on a log scale. The asterisk denotes statistical significance (p<0.05).

	pLuminex		mLuminex		MSD			
Assay	**%**	**n/N**	**%**	**n/N**	**%**	**n/N**	**p-value**	
IL-1β	3.7	1/27	0.0	0/31	**29.0**	9/31	0.003	**
IL-6	96.7	29/30	**96.8**	30/31	**96.8**	30/31		
IL-8	88.5	23/26	96.8	30/31	**100.0**	31/31	0.097	
TNF-α	6.7	2/30	0.0	0/31	**87.1**	27/31	<0.001	**
IL-10	0.0	0/30	0.0	0/31				
IL-13	0.0	0/25	0.0	0/31				
IL-12	100.0	27/27	100.0	31/31				
RANTES	**80.7**	25/31	25.8	8/31			<0.001	**
Eotaxin	95.7	22/23	96.8	30/31				
IL-17	0.0	0/24	0.0	0/31				
MIP-1α	13.3	4/30	**45.2**	16/31			0.004	**
GM-CSF	0.0	0/28	0.0	0/31				
MIP-1β	**63.3**	19/30	6.45	2/31			<0.001	**
MCP-1	100.0	19/19	100.0	31/31				
IL-15	25.0	7/28	**100.0**	31/31			<0.001	**
IL-5	0.0	0/27	0.0	0/31				
IFN-γ	0.0	0/28	0.0	0/31				
IFN-α	3.7	1/27	**16.1**	5/31			0.625	
IL-1RA	**38.7**	12/31	32.26	10/31			0.727	
IL-2	0.0	0/29	0.0	0/31				
IL-7	**4.6**	1/22	0.0	0/31			1.000	
IP-10	100.0	20/20	100.0	31/31				
IL-2R	**55.2**	16/29	25.8	8/31			0.034	*
MIG	**23.8**	6/26	9.8	3/31			0.025	*
IL-4	0.0	0/31	0.0	0/31				

Table 4.7 – Proportions of SF samples (N=31) above the limit of quantification (LOQ) for each analyte and platform given in Table 4.6**.** The highest proportion for each analyte is given in bold. Matched proportions were compared by McNemar's exact test or Cochran's Q test. Asterisks (*) denote significance (p<0.05) and a double asterisk (**) denotes significance (q<0.05) after correction for multiple comparisons by Holm's method.

On comparison of the Luminex bead systems, 8 analytes [IL-10, IL-13, IL-17, GM-CSF, IL-5, IFN-γ, IL-2 & IL-4] were <LOQ in all samples tested, and 3 [IL-12, MCP-1 & IP-10] were >LOQ in all samples on both Luminex platforms [Table 4.7]. Eotaxin was >LOQ in all but one (same) sample on both systems. Reflecting the lower LOQs, RANTES, MIP-1β, IL-2R and MIG were >LOQ in significantly more SF samples on the pLuminex platform. However, MIP-1α and IL-15 were quantifiable in significantly more samples using the mLuminex platform even though the LOQs were equivalent on both platforms.

4.4.4 Spike Recovery

The spike recoveries for the MSD 4-plex and mLuminex 25-plex are given in Table 4.8 and Table 4.9 respectively. For the MSD assays all spike recoveries were satisfactory except for lower spike of IL-6 (76.6%), which was <5% outside of the acceptable range. For the same 4 analytes on the mLuminex platform, IL-6 and IL-8 had acceptable spike recoveries for both spike concentrations. Recoveries at both spike concentrations were below acceptable for IL-1β and TNF-α. However, they were <5% outside the acceptable range at the higher spike and <10% outside at the lower spike.

The majority of the remaining mLuminex analytes (14 out 21) had an acceptable spike recovery at the higher spike-A concentrations. Four analytes had recoveries <5% outside of the acceptable threshold and one was <10% outside of satisfactory. At the lower spike-B concentrations, 8 of the remaining analytes had acceptable spike recoveries with a further 5 analytes just falling short of satisfactory; 2 were <5% outside and 3 were <10% outside of the acceptable range. The spike recovery of MIG at both concentrations was in excess of the acceptable threshold. There was a potentially spurious negative (-88.4%) recovery of MCP-1 for spike-B.

The median difference in %recovery between high and low spike concentrations was 14.4% (IQR 5.9 – 23.3) for MSD and 10.1% (IQR 2.6 – 14.6) for mLuminex.

		Spike-A			Spike-B		
Assay	**C_U**	**C_{Spike}**	**C_{Obs}**	**Recovery (%)**	**C_{Spike}**	**C_{Obs}**	**Recovery (%)**
IL-1β	0.6	833.3	905.2	108.6	208.3	201.3	96.3
IL-6	221.4	833.3	1072.6	102.2	208.3	381.0	76.6
IL-8	64.5	833.3	940.4	105.1	208.3	249.0	88.6
TNF-α	2.2	833.3	791.2	94.7	208.3	191.7	91.0

Table 4.8 – Spike recoveries for MSD 4-plex assays. C_U is the measured concentration in the unspiked SF sample corrected for dilution by the spike. C_{Obs} is the measured concentration in the spiked sample. C_{Spike} is the expected spike concentration. All concentrations are in pg/ml. All measured concentrations were above LOQ. Green highlighting indicates an acceptable recovery of 100% ±20%.

		Spike-A			Spike-B		
Assay	**C_U**	**C_{Spike}**	**C_{Obs}**	**Recovery (%)**	**C_{Spike}**	**C_{Obs}**	**Recovery (%)**
IL-6	538.2	345.8	859.8	93.0	38.4	569.1	80.4
IL-8	248.7	1066.7	1157.1	85.2	118.5	343.5	80.0
IL-12	28.2	541.7	397.6	68.2	60.2	63.1	58.1
Eotaxin	9.5	241.7	201.4	79.4	26.9	31.4	81.4
MIP-1α	17.9	2083.3	1719.9	81.7	231.5	201.8	79.4
MCP-1	1268.9	508.3	1708.3	86.4	56.5	1219.0	-88.4
IL-15	67.2	2025.0	1382.5	65.0	225.0	243.1	78.2
IFN-α	34.9	1075.0	968.2	86.8	119.4	109.3	62.3
IL-1RA	67.9	2333.3	1905.3	78.7	259.3	242.5	67.3
IP-10	22.6	75.0	81.7	78.7	8.3	30.0	89.2
IL-2R	65.5	1750.0	1705.2	93.7	194.4	222.4	80.7
MIP-1β	14.8	600.0	500.3	80.9	66.7	70.0	82.7
IL-5	1.1	194.2	178.8	91.5	21.6	19.7	86.3
IL-1β	1.7	691.7	527.2	76.0	76.9	55.1	70.0
TNF-α	0.8	725.0	561.1	77.3	80.6	59.6	72.9
IL-10	0.8	1825.0	1631.8	89.4	202.8	187.4	92.0
IL-13	3.0	1691.7	1601.1	94.5	188.0	185.7	97.1
IL-17	6.0	3233.3	2779.2	85.8	359.3	263.4	71.7
GM-CSF	3.1	1445.8	1023.7	70.6	160.6	113.3	68.6
IFN-γ	0.9	875.0	715.4	81.7	97.2	85.5	87.1
IL-2	0.5	100.0	83.5	83.0	11.1	8.0	68.0
IL-4	2.6	3550.0	3043.5	85.7	394.4	343.8	86.5
RANTES	2.8	291.7	289.3	97.9	32.4	16.3	41.7
IL-7	8.4	816.7	699.9	84.7	90.7	71.2	69.3
MIG	4.0	200.0	356.3	176.2	22.2	39.7	160.7

Table 4.9 – Spike recoveries for mLuminex 25-plex assays. C_U is the measured concentration in the unspiked SF sample corrected for dilution by the spike. C_{Obs} is the measured concentration in the spiked sample. C_{Spike} is the expected spike concentration. All concentrations are in pg/ml. Red text indicates measured concentrations below LOD that were substitution with LOD/2. Blue text indicates measured concentrations below LOQ that were substitution with LOQ/2. Green highlighting indicates an acceptable recovery of 100% ±20%.

4.4.5 Intra-assay Precision

The coefficient of variation (CV) for replicate measurements was significantly higher for IL-6 and IL-8 on the pLuminex platform than mLuminex and MSD platforms between which there was no difference [Table 4.10]: IL6 $p<0.001$ for pLuminex vs. mLuminex and pLuminex vs. MSD, and $p>0.999$ for mLuminex vs. MSD; IL8 $p=0.017$ for pLuminex vs. mLuminex, $p=0.080$ for pLuminex vs. MSD, and $p>0.999$ for mLuminex vs. MSD. The replicate CV was also significantly greater on the pLuminex platform for IL-12 and IP-10 before correction for multiple comparisons. However, the median replicate CV for all analytes was <20% irrespective of platform. Similarly, there were no differences between platforms in the number of samples with replicate CVs $\leq$20% for any of the analytes [Table 4.11].

4.4.6 ELISA Validation

All SF samples tested by the TNF-α ELISA were below LOD (6.52pg/ml) and so could not be used to validate Luminex and MSD measurements. However, all samples tested by the IL-6 ELISA were >LOQ (9.38 pg/ml).

Table 4.12 shows the concordance and intraclass correlation coefficients. The concordance of MSD and ELISA measurements was moderate (CCC=0.901, 95%CI 0.816 – 0.960). Both Luminex platforms had a poor concordance (CCC<0.90) with ELISA measurements. However, the 95%CIs of the CCCs for all three platforms overlapped suggesting no platforms differences for concordance with ELISA measurements. The ICCs for absolute agreement with ELISA were not reliable because the lower bounds of the 95%CIs are negative. However, the ICC for consistency of measurements between by all three platforms with ELISA was >0.90 with the lower bound of the 95%CI >0.80. Therefore all three platforms had “almost perfect” relative agreement with ELISA measurements.

Weighted Deming regression [Figure 4.4] for IL-6 measurements by the MSD and mLuminex platforms showed no systematic or proportional bias compared to ELISA since the 95%CI for the intercept included 0 and the 95%CI for the slope included

	pLuminex		mLuminex		MSD			
Assay	**CV**	**IQR**	**CV**	**IQR**	**CV**	**IQR**	**p-value**	
IL-6	8.2	4.8 - 12.6	2.4	0.9-4,4	3.69	2.02-8.80	<0.001	**
IL-8	7.7	3.6 - 12.6	2.1	1.1-6.2	4.21	1.51-8.35	0.0135	**
TNF-a	12.6	2.6 - 28.6			9.66	4.90 -14.95	0.7221	
IL-12	10.8	5.3 -17.2	4.8	1.7-9.2			0.0186	*
RANTES	8.1	6.2 - 17.1	4.3	2.1 - 8.5			0.0627	
Eotaxin	14.1	5.5 - 21.4	8.9	3.92-21.17			0.0693	
MIP-1a	13.3	9.7 - 18.2	10.9	2.00-12.58			0.1141	
MIP-1b	10.3	6.9 - 15.5	11.1	5.36-16.56			0.7239	
MCP-1	5.4	3.0 - 9.9	5.8	1.49-23.4			0.1590	
IL-15	12.9	8.9 - 19.1	8.8	3.99-17.37			0.0651	
IL-1RA	8.8	0.0 - 28.4	18.3	10.53-29.26			0.2970	
IP-10	6.1	4.1 - 10.9	4.0	1.43-6.81			0.0449	*

Table 4.10 – Average replicate CVs as a measure of intra-assay precision. The median replicate CV and interquartile range (IQR) are given for analytes with at least 10 matched samples >LOD on each platform under comparison. There were no samples >LOD for TNF-α on the mLuminex platform. The average CV for each analyte was compared across platforms, using the Kruskal-Wallis test for IL-6 & IL-8, and Mann-Whitney U-test for the remaining analytes. Asterisks (*) denote significance (p<0.05) and a double asterisk (**) denotes significance (q<0.05) after correction for multiple comparisons by Holm's method.

	pLuminex		mLuminex		MSD		
Assay	**%**	**n/N**	**%**	**n/N**	**%**	**n/N**	**p-value**
IL-6	86.2	25/29	96.7	29/30	100.0	30/30	0.1353
IL-8	92.3	24/26	93.5	29/31	100.0	31/31	0.3679
TNF-a	61.5	8/23			77.4	24/31	0.6250
IL-12	81.5	22/27	90.3	28/31			0,7266
RANTES	84.6	22/26	100.0	14/14			0.6250
Eotaxin	69.6	26/23	74.2	23/31			0.7539
MIP-1a	83.3	10/12	93.8	15/16			1.000
MIP-1b	92.9	26/28	91.7	11/12			1.000
MCP-1	94.7	18/19	74.2	23/31			1.000
IL-15	78.6	22/28	80.6	25/31			1.000
IL-1RA	59.1	13/22	57.9	11/19			1.000
IP-10	93.3	28/30	93.5	29/31			1.000

Table 4.11 – Proportion of replicate CVs <20%. The percentage of replicate CVs below the accepted threshold of 20% are given for analytes with at least 10 matched samples >LOD on each platform under comparison. There were no samples >LOD for TNF-α on the mLuminex platform. Matched proportions were compared by McNemar's exact test or Cochran's Q test (for IL-6 & IL-8). Asterisks (*) denote significance (p<0.05) and a double asterisk (**) denotes significance (q<0.05) after correction for multiple comparisons by Holm's method.

	Concordance		ICC Absolute		ICC Consistency		
	CCC	**95%CI**	**ICC**	**95%CI**	**ICC**	**95%CI**	**N**
pLuminex	0.768	0.481 – 0.906	0.786	-0.059 – 0.960	0.956	0.833 – 0.988	10
mLuminex	0.805	0.503 – 0.932	0.825	-0.050 – 0.972	0.969	0.852 – 0.994	8
MSD	0.902	0.816 – 0.960	0.916	-0.010 – 0.988	0.993	0.970 – 0.988	11

Table 4.12 – Concordance correlation coefficient (CCC) and intraclass correlation coefficients (ICC) for absolute agreement and consistency for ELISA and multiplex measurements of IL-6. Coefficients, 95%CIs and the number of matched samples (N) used for calculations are given for each platform compared to ELISA.

1. Arguably the slope for mLuminex (1.45, 95%CI 0.99 – 1.91) comes very close to qualifying for a proportional bias due to the lower bound of the 95%CI being close to 1. Although there was no systematic bias for pLuminex measurements (95%CI of intercept included zero), there was a proportional bias with a regression coefficient (slope) of 1.84 (95%CI 1.68 – 2.00).

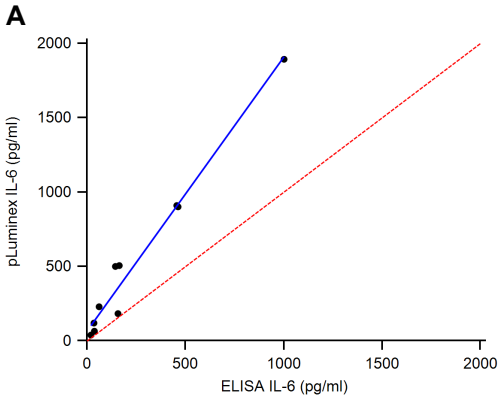
Bland-Altman analysis of IL-6 measurements by all 3 platforms vs. ELISA indicated a proportional bias since there was a significant linear relationship between the difference and average measurements [Figure 4.5]. This is confirmed by the ratio vs. difference plots for all three comparisons that show no correlation and a mean (ratio) bias that does not equal 1 (i.e. 95%CI doesn't include 1).

4.4.7 Cross-platform Agreement

There were 6 analytes (IL-6, IL-8, IL-12, Eotaxin, MCP-1 and IP-10) that were >LOQ in at least 15 matched SF samples tested on each platform to allow assessment of cross-platform agreement.

Intraclass correlation and concordance

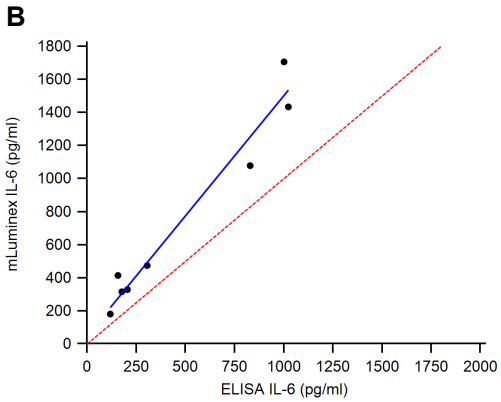
Table 4.13 and Table 4.14 show the intraclass and concordance correlation coefficients for cross-platform agreement. The absolute ICC demonstrated "almost perfect" agreement for IL-6 and IL-8 measurements by the p- and m-Luminex platform (ICC>0.9 for both cytokines). This is reflected by a moderate concordance for IL-6 (CCC=0.916



Method	Pooled CV (%)	N=10
pLuminex	13.7	
ELISA	6.2	

Regression Parameters

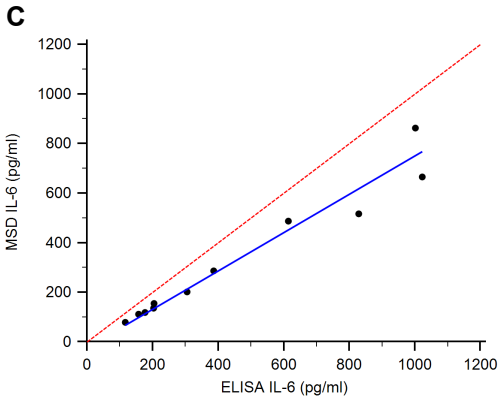
	Coeff.	SE	95%CI
Intercept	68.43	43.95	-30.99 to 167.84
Slope	1.84	0.07	1.68 to 2.00



Method	Pooled CV (%)	N=8
mLuminex	5.9	
ELISA	6.3	

Regression Parameters

	Coeff.	SE	95%CI
Intercept	50.60	60.60	-92.71 to 193.91
Slope	1.45	0.196	0.99 to 1.91



Method	Pooled CV (%)	N=11
MSD	6.4	
ELISA	9.7	

Regression Parameters

	Coeff.	SE	95%CI
Intercept	--18.67	32.36	-90.79 to 53.44
Slope	0.76	0.12	0.50 to 1.03

Figure 4.4 – Weighted Deming regression of IL-6 measurements by the [A] pLuminex, [B] mLuminex and [C] MSD platform compared to ELISA. *Identity* (slope=1) is represented by the red hashed line. The pooled coefficients of variation (CV) for measurements by each platform used for weighting the Deming regression are indicated. The regression parameters and their 95% confidence intervals are shown for each comparison. The intercept indicates systematic bias (error) and the slope represents proportional bias.

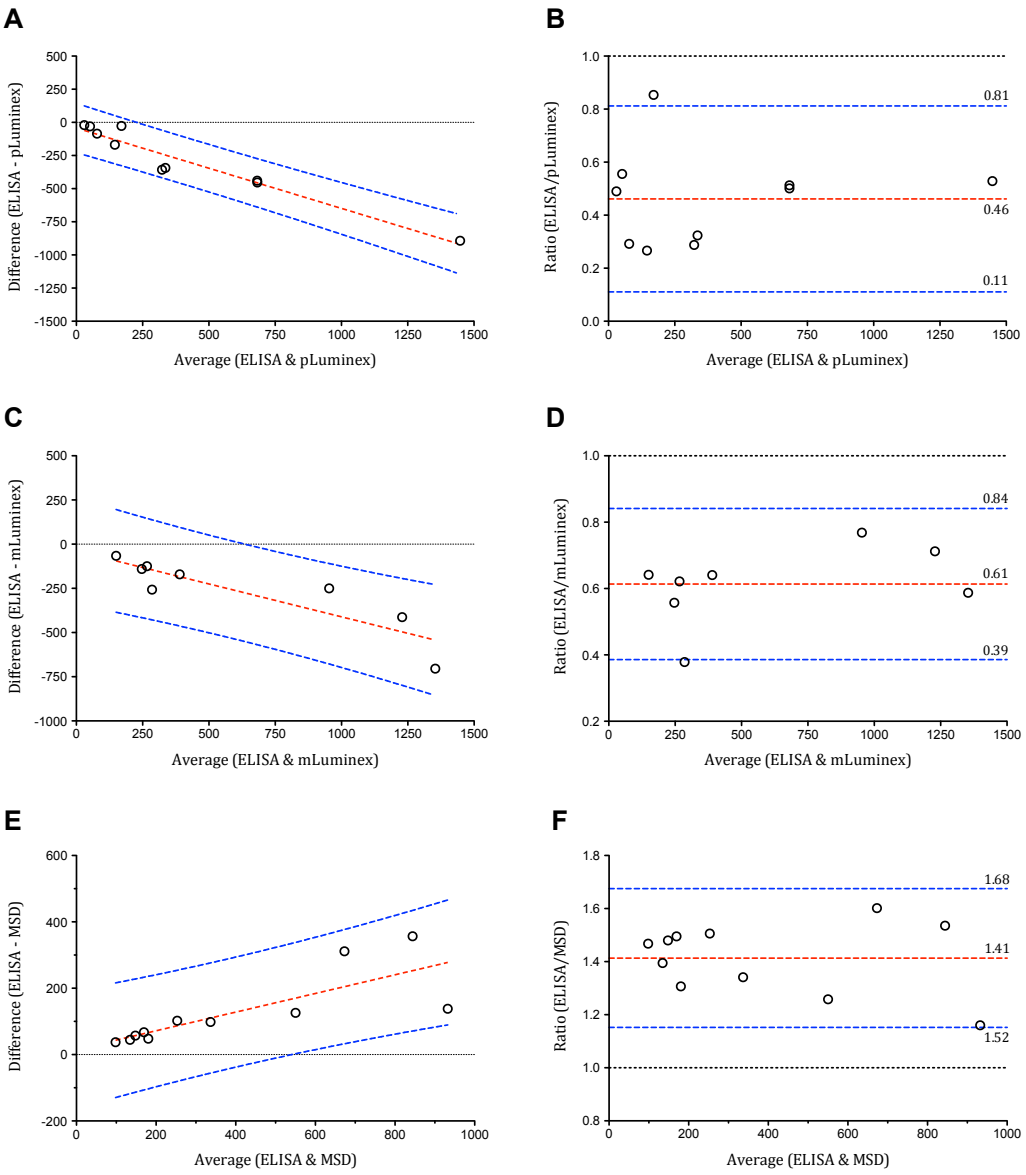


Figure 4.5 – Bland-Altman plots for agreement of IL-6 measurements by the [A & B] pLuminex, [C & D] mLuminex and [E & F] MSD platforms with ELISA. The red-hashed line represents the error (bias) and the blue-hashed line represents the 95% limits of agreement (LOA). **[A, C & E]** Classical plots for the difference in measurements vs. the average of measurements. The hashed line at y=0 represents zero systematic bias. When there was a significant relationship between difference and average measurements i.e. proportional bias a regression line (red dotted line) and 95% prediction bands (blue dotted lines) were drawn to represent the mean bias and 95% limits of agreement. **[B, D & F]** Plots for the ratio of measurements vs. the average of measurements. The hashed line at y=1 represents zero proportional bias. The mean proportional bias and 95% limits of proportional agreement are shown. There was a constant proportional bias for each platform and the LOA do not include 1.

		Absolute		Consistency		
Assay	**Platforms**	**ICC**	**95%CI**	**ICC**	**95%CI**	**N**
IL-6	3 platforms	0.767	0.124 – 0.926	0.962	0.932 – 0.981	29
	pLum – mLum	0.918	0.136 – 0.979	0.974	0.945 – 0.988	29
	pLum – MSD	0.797	-0.054 – 0.949	0.956	0.908 – 0.979	29
	mLum – MSD	0.632	-0.034 – 0.900	0.957	0.911 – 0.979	30
IL-8	3 platforms	0.651	0.057 – 0.883	0.945	0.895 – 0.975	23
	pLum – mLum	0.966	0.883 – 0.988	0.975	0.942 – 0.989	23
	pLum – MSD	0.587	-0.056 – 0.881	0.918	0.818 – 0.965	23
	mLum – MSD	0.560	-0.290 – 0.871	0.952	0.902 – 0.977	30
IL-12	pLum – mLum	0.472	-0.054 – 0.820	0.884	0.761 – 0.945	27
Eotaxin	pLum – mLum	0.177	-0.144 – 0.512	0.237	-0.206 – 0.600	21
MCP-1	pLum – mLum	0.728	-0.029 – 0.920	0.877	0.709 – 0.951	19
IP-10	pLum – mLum	0.548	-0.095 – 0.836	0.800	0.621 – 0.899	30

Table 4.13 – Intraclass correlation coefficients (ICC) for absolute agreement & consistency. The ICC and 95%CI are given for each analyte and platform comparison. The number of matched samples (N) used for ICC calculations is stated.

95%CI 0.867-0.965) and substantial concordance for IL-8 (CCC=0.965 95%CI 0.938-0.991) measurements between Luminex platforms. The ICCs for absolute agreement between MSD and either Luminex platform were not reliable because the lower bounds of the 95%CIs are negative. Likewise, there was a poor concordance (CCC<0.90) for IL-6 and IL-8 measurements between MSD and either Luminex platform. The ICC for absolute agreement between Luminex platforms for IL-12, MCP-1 and IP-10 was unreliable (negative lower bounds of the 95%CIs) and the concordances were also poor. The relative agreement indicated by ICC for consistency between platforms for all analytes except Eotaxin was "almost perfect" (ICC>0.80). Agreement between Luminex platforms for Eotaxin could not be assessed as the ICCs and CCC had negative lower bounds of the 95%CIs.

Weighted Deming regression

IL-6 [Figure 4.6 & C] and IL-8 [Figure 4.7B & C] measurements by the MSD platform demonstrated proportional bias in agreement (95%CI of slopes did not include 1) with no systematic bias (95%CI of intercepts include 0) when compared to either Luminex platform. IL-8 measurements by the pLuminex and mLuminex platforms were close to

Assay	Platforms	CCC	95%CI	N
IL-6	pLum – mLum	0.916	0.867 – 0.965	29
	pLum – MSD	0.792	0.688 – 0.895	29
	mLum – MSD	0.624	0.491 – 0.757	30
IL-8	pLum – mLum	0.965	0.938 – 0.991	23
	pLum – MSD	0.576	0.407 – 0.746	23
	mLum – MSD	0.533	0.412 – 0.693	30
IL-12	pLum – mLum	0.462	0.306 – 0.618	27
Eotaxin	pLum – mLum	0.170	-0.137 – 0.477	21
MCP-1	pLum – mLum	0.717	0.537 – 0.897	19
IP-10	pLum – mLum	0.540	0.377 – 0.702	30

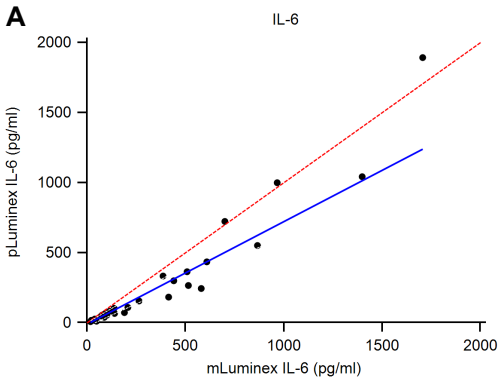
Table 4.14 – Concordance correlation coefficients (CCC). The CC and 95%CI are given for each analyte and pairwise platform comparison. The number of matched samples (N) used for CCC calculations is stated.

identical [Figure 4.7A], with the slope and intercept not significantly different from 1 and zero respectively. However, IL-6 measurements by the Luminex platforms showed a small proportional bias with a slope of 0.7 (95%CI 0.55-0.86) [Figure 4.7A].

pLuminex and mLuminex measurements of IL-12 [Figure 4.8A], MCP-1 [Figure 4.9A] and IP-10 [Figure 4.9B] demonstrated proportional errors in agreement with no systematic bias. Eotaxin measurements by Luminex platforms showed poor correlation (r=0.30, p=0.19) with a wide dispersal of paired measurements [Figure 4.8B]. Therefore further comment on regression parameters is not made.

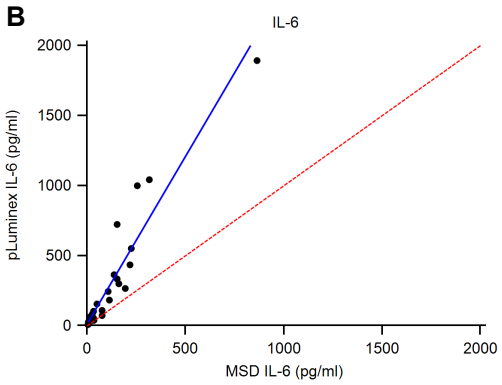
Bland-Altman Analysis

Comparison of MSD with either Luminex platforms demonstrated a proportional bias for IL-6 [Figure 4.10 C & D, E & F] and IL-8 measurements [Figure 4.11 C & D, E & F]. This is indicated by a significant linear relationship between the difference and average measurements, but no correlation between the ratio vs. difference and a mean (ratio) bias that does not equal 1. There is good agreement between Luminex platforms for IL-8, with a mean systematic bias of 12.1pg/ml (95%CI -40.36 to 16.23pg/ml) and no proportional bias [Figure 4.11 A & B]. At first glance, there appeared to be good agreement between Luminex platforms for IL-6 [Figure 4.10 A & B], but exclusion of 3 outliers resulted in features of proportional bias with no systematic error.



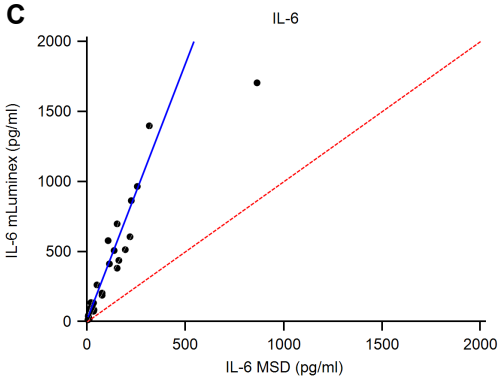
Method	Pooled CV (%)	N=29
mLuminex	8.0	
pLuminex	19.5	

Regression Parameters	Coeff.	SE	95%CI
Intercept	-19.0	15.5	-50.98 to 13.02
Slope	0.70	0.08	0.55 to 0.86



Method	Pooled CV (%)	N=23
MSD	9.1	
pLuminex	19.5	

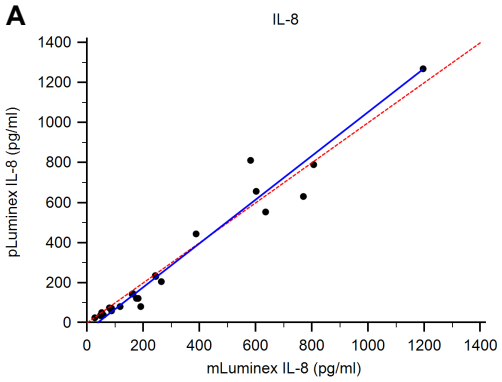
Regression Parameters	Coeff.	SE	95%CI
Intercept	12.12	51.01	-92.38 to 116.61
Slope	2.39	0.64	1.09 to 3.69



Method	Pooled CV (%)	N=30
MSD	9.1	
mLuminex	7.9	

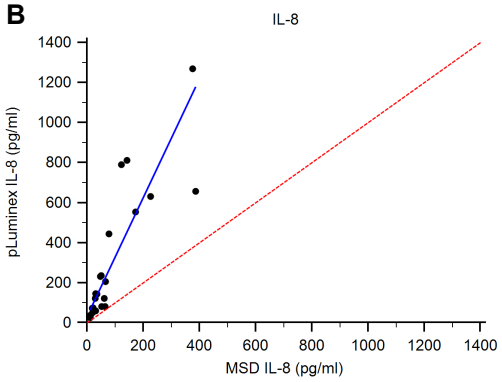
Regression Parameters	Coeff.	SE	95%CI
Intercept	-18.19	23.94	-67.25 to 30.86
Slope	3.87	0.38	3.09 to 4.64

Figure 4.6 – Weighted Deming regression for comparison of IL-6 measurements. [A] pLuminex vs. mLuminex, **[B]** pLuminex vs. MSD and **[C]** mLuminex vs. MSD. The dotted red line is the line of identity (y=x). Regression parameters (slope and intercept) are presented with their 95%CIs. If the 95%CI of the intercept includes 0, there is no absolute (systematic) bias; if the 95%CI of the slope includes 1, there is no proportional bias.



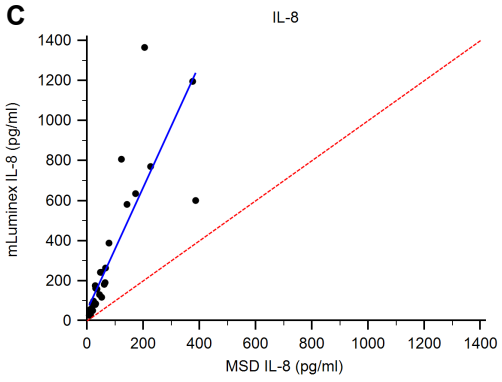
Method	Pooled CV (%)	N=23
mLuminex	22.5	
pLuminex	11.2	

Regression Parameters	Coeff.	SE	95%CI
Intercept	-40.83	11.48	-64.65 to -17.02
Slope	1.09	0.06	0.96 to 1.23



Method	Pooled CV (%)	N=23
MSD	9.2	
pLuminex	11.2	

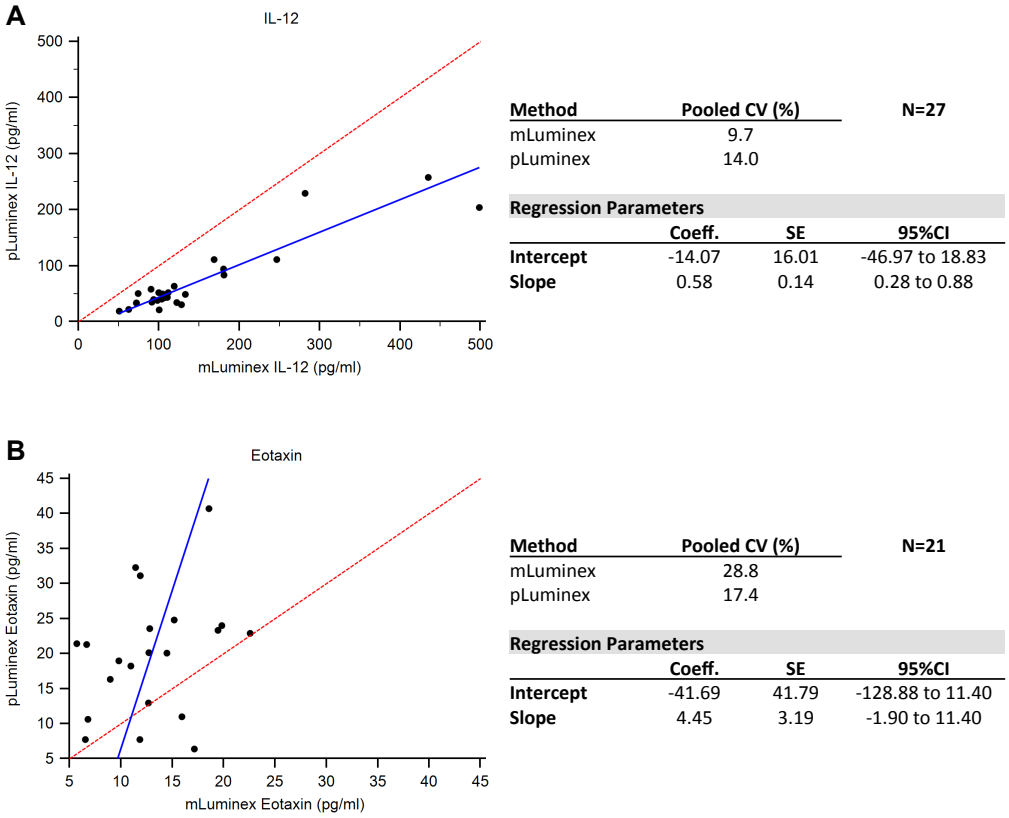
Regression Parameters	Coeff.	SE	95%CI
Intercept	34.02	42.76	-52.58 to 120.63
Slope	2.97	0.81	1.29 to 4.65



Method	Pooled CV (%)	N=30
MSD	10.0	
mLuminex	20.9	

Regression Parameters	Coeff.	SE	95%CI
Intercept	52.78	43.56	-39.30 to 141.87
Slope	3.07	0.99	1.05 to 5.10

Figure 4.7 – Weighted Deming regression for comparison of IL-8 measurements. [A] pLuminex vs. mLuminex, **[B]** pLuminex vs. MSD and **[C]** mLuminex vs. MSD. The dotted red line is the line of identity (y=x). Regression parameters (slope and intercept) are presented with their 95%CIs. If the 95%CI of the intercept includes 0, there is no absolute (systematic) bias; if the 95%CI of the slope includes 1, there is no proportional bias.



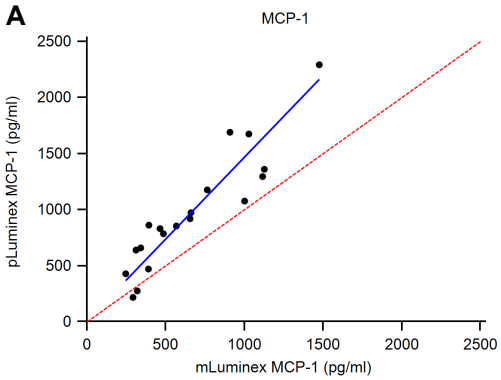
Method	Pooled CV (%)	N=27
mLuminex	9.7	
pLuminex	14.0	

Regression Parameters	Coeff.	SE	95%CI
Intercept	-14.07	16.01	-46.97 to 18.83
Slope	0.58	0.14	0.28 to 0.88

Method	Pooled CV (%)	N=21
mLuminex	28.8	
pLuminex	17.4	

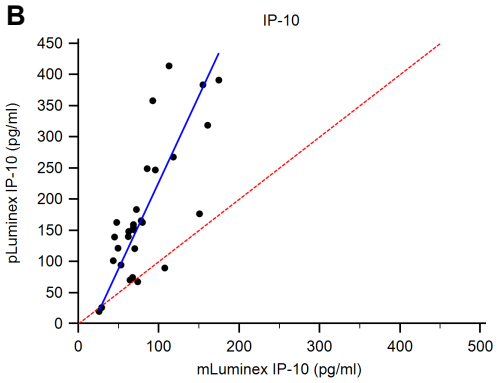
Regression Parameters	Coeff.	SE	95%CI
Intercept	-41.69	41.79	-128.88 to 11.40
Slope	4.45	3.19	-1.90 to 11.40

Figure 4.8 – Weighted Deming regression for comparison of [A] IL-12 and [B] Eotaxin measurements between Luminex platforms. The dotted red line is the line of identity (y=x). Regression parameters (slope and intercept) are presented with their 95%CIs. If the 95%CI of the intercept includes 0, there is no absolute (systematic) bias; if the 95%CI of the slope includes 1, there is no proportional bias.



Method	Pooled CV (%)	N=19
mLuminex	11.7	
pLuminex	13.2	

Regression Parameters	Coeff.	SE	95%CI
Intercept	12.50	100.81	-199.31 to 224.30
Slope	1.46	0.16	1.11 to 1.80



Method	Pooled CV (%)	N=30
mLuminex	11.0	
pLuminex	14.5	

Regression Parameters	Coeff.	SE	95%CI
Intercept	--54.26	14.65	-83.99 to -24.52
Slope	2.84	0.088	2.66 to 3.02

Figure 4.9 – Weighted Deming regression for comparison of [A] MCP-1 and [B] IP-10 measurements between Luminex platforms. The dotted red line is the line of identity (y=x). Regression parameters (slope and intercept) are presented with their 95%CIs. If the 95%CI of the intercept includes 0, there is no absolute (systematic) bias; if the 95%CI of the slope includes 1, there is no proportional bias.

pLuminex and mLuminex measurements of IL-12 [Figure 4.12 A & B], Eotaxin [Figure 4.12 C & D] and IP-10 [Figure 4.13 C & D] demonstrated proportional errors in agreement. MCP-1 showed a significant ($p=0.040$) linear relationship between the difference and average measurements [Figure 4.13A], but not for ratio vs. difference [Figure 4.13B]. However, the mean (ratio) bias was not significantly different to 1 (ratio 1.76, 95%CI 0.86 – 2.67) raising doubt over a proportional bias.

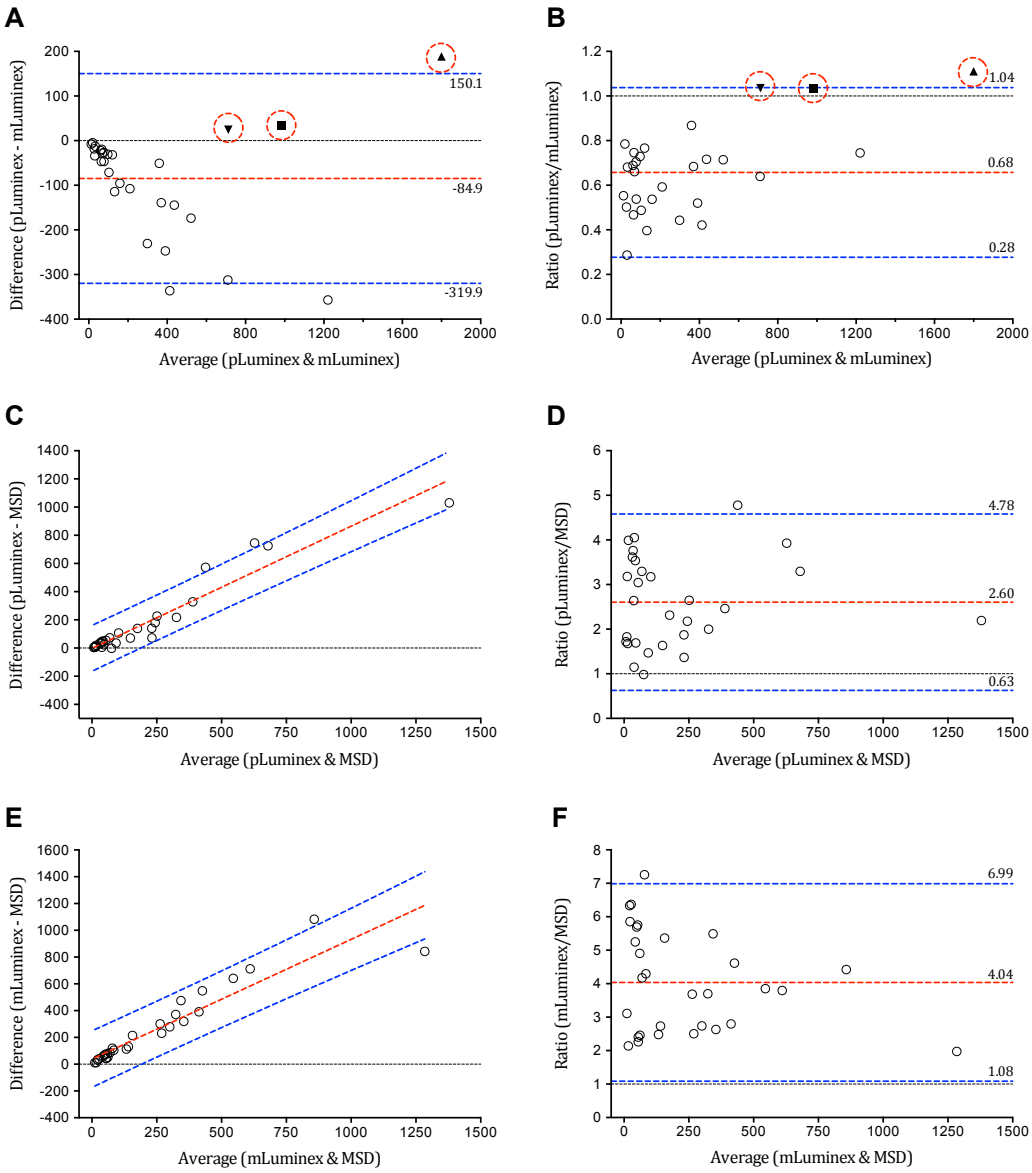


Figure 4.10 – Bland-Altman plots for agreement of IL-6 measurements between multiplex platforms. [A & B] pLuminex vs. mLuminex. **[C & D]** pLuminex vs. MSD. **[E & F]** mLuminex vs. MSD. The left-hand plots **[A, C & E]** are classical plots of difference vs. average measurements. The black-hashed line at y=0 represents zero absolute (systematic) bias. The mean bias (red-hashed line) and 95% limits of agreement (blue-hashed line) are shown. When there was a significant relationship between difference and average measurements i.e. proportional bias **[C & E]** a regression line (red dotted line) and 95% prediction bands (blue dotted lines) were drawn to represent the mean bias and 95% limits of agreement. Outliers are highlighted in red circles. The right-hand plots **[B, D & F]** are of ratio vs. average measurements. The black line represents no proportional bias (ratio=1). The mean proportional bias and 95% limits of proportional agreement are shown. 95% confidence intervals or bands are not shown in any graph.

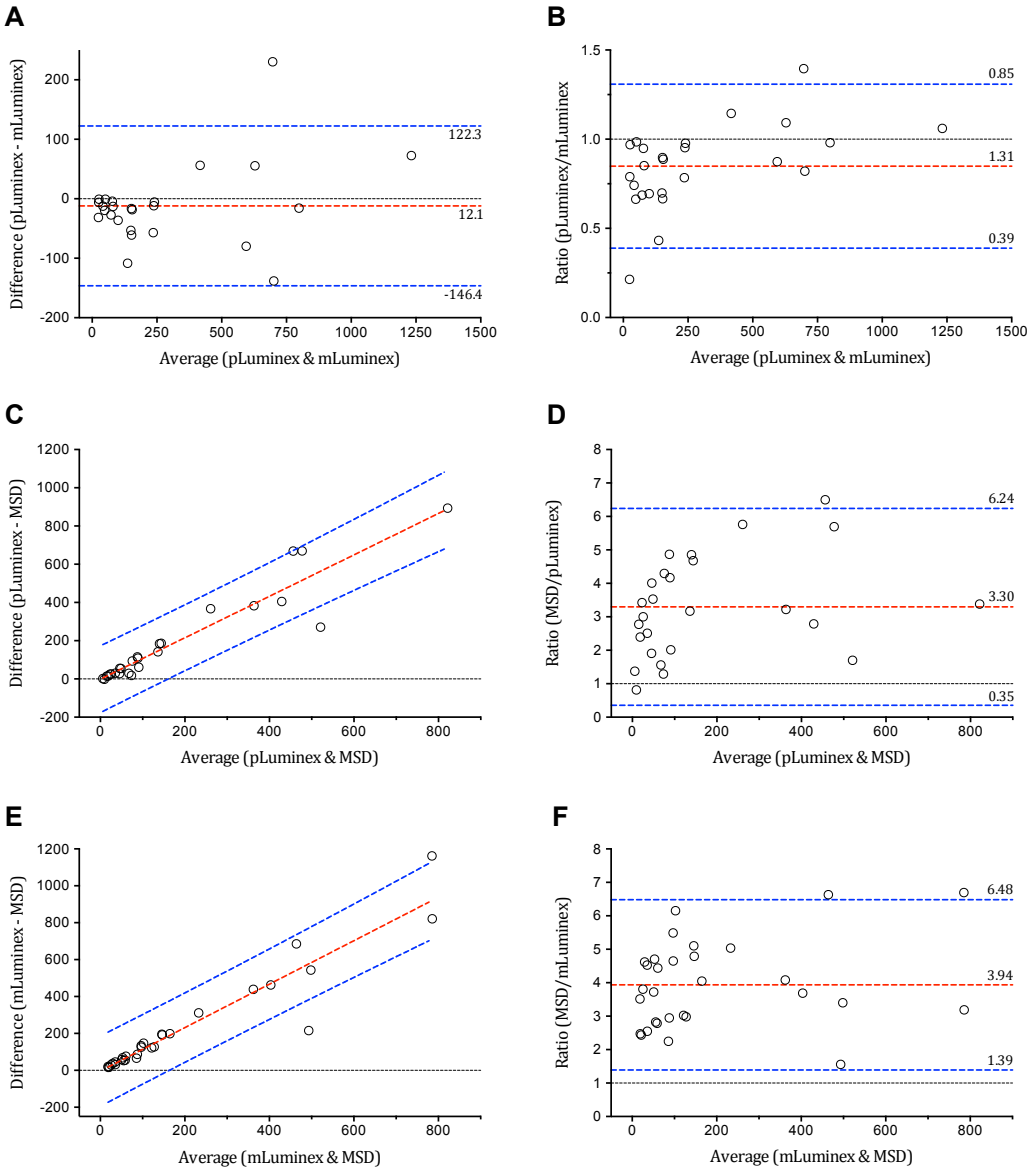


Figure 4.11 – Bland-Altman plots for agreement of IL-8 measurements between multiplex platforms. [A & B] pLuminex vs. mLuminex. **[C & D]** pLuminex vs. MSD. **[E & F]** mLuminex vs. MSD. The left-hand plots **[A, C & E]** are of difference vs. average measurements. The black line represents zero absolute (systematic) bias. The mean bias (red hashed line) and 95% limits of agreement (blue hashed line) are shown. When there was a significant relationship between difference and average measurements i.e. proportional bias **[C & E]** a regression line (red dotted line) and 95% prediction bands (blue dotted lines) were drawn to represent the mean bias and 95% limits of agreement. The right-hand plots **[B, D & F]** are of ratio vs. average measurements. The black line represents no proportional bias (ratio=1). The mean proportional bias and 95% limits of proportional agreement are shown. 95% confidence intervals or bands are not shown in any graph.

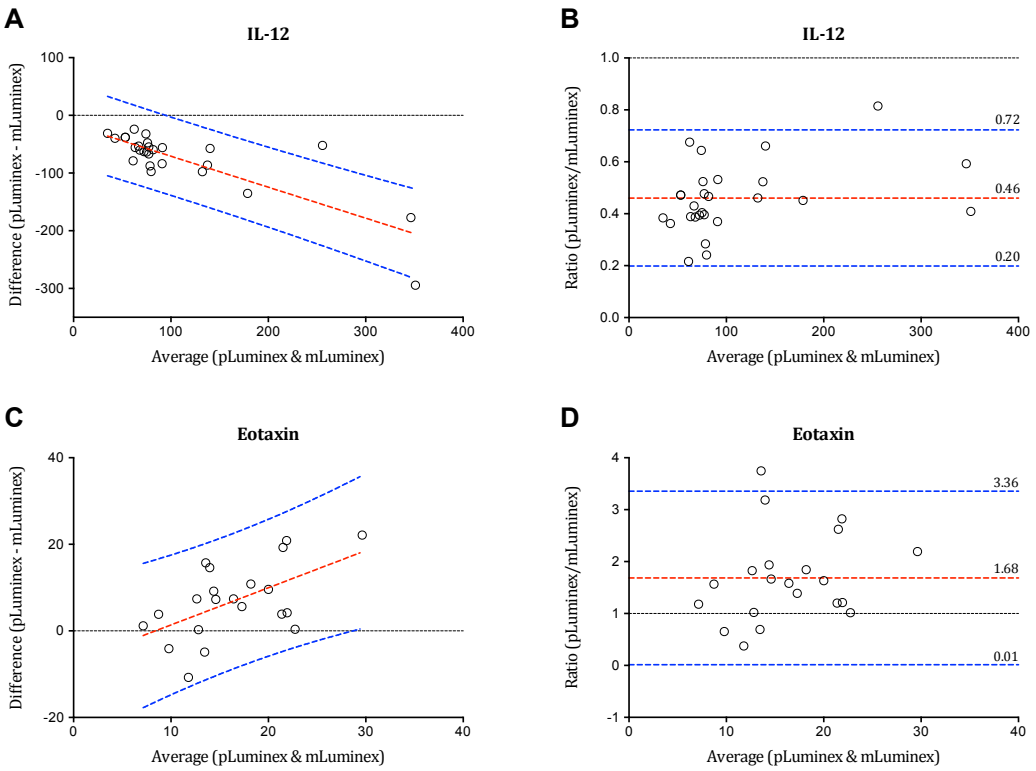


Figure 4.12 – Bland-Altman plots for agreement of [A & B] IL-12 and [C & D] Eotaxin measurements between Luminex platforms. The left-hand plots **[A & C]** are of difference vs. average measurements. The black line represents zero absolute (systematic) bias. There was a significant relationship between difference and average measurements i.e. proportional bias, so a regression line (red dotted line) and 95% prediction bands (blue dotted lines) were drawn to represent the mean bias and 95% limits of agreement. The right-hand plots **[B & D]** are of ratio vs. average measurements. The black line represents no proportional bias (ratio=1). The mean proportional bias and 95% limits of proportional agreement are shown. 95% confidence intervals or bands are not shown in any graph.

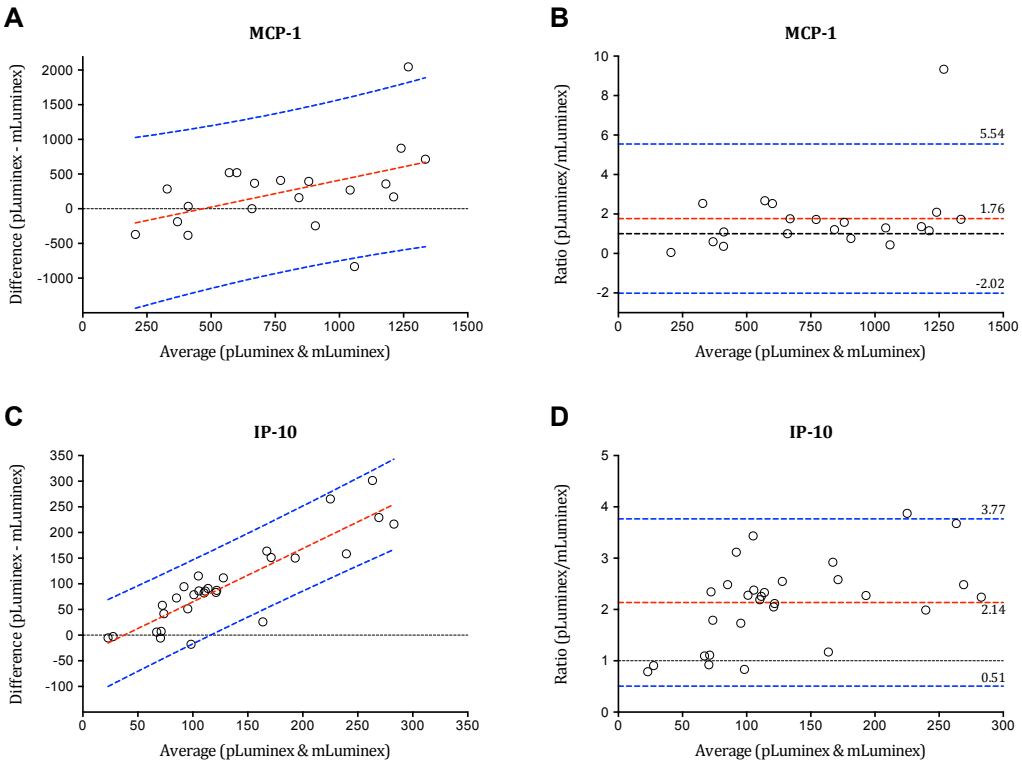


Figure 4.13 – Bland-Altman plots for agreement of [A & B] MCP-1 and [C & D] IP-10 measurements between Luminex platforms. The left-hand plots **[A & C]** are of difference vs. average measurements. The black line represents zero absolute (systematic) bias. There was a significant relationship between difference and average measurements i.e. proportional bias, so a regression line (red dotted line) and 95% prediction bands (blue dotted lines) were drawn to represent the mean bias and 95% limits of agreement. The right-hand plots **[B & D]** are of ratio vs. average measurements. The black line represents no proportional bias (ratio=1). The mean proportional bias and 95% limits of proportional agreement are shown. 95% confidence intervals or bands are not shown in any graph.

4.5 Discussion

This chapter provides a comprehensive comparison of two prominent multiplex immunoassay technologies, Luminex and MSD, in the context of OA biomarker research using SF samples from real patients. Furthermore, both polystyrene- and magnetic-bead systems of the Luminex platform were considered. Taken together with the findings in Chapter 3, all 3 platforms are (partially) validated for the analysis of SF. However, the comparison highlights important differences in assay performance across a range of parameters, and poor agreement of analyte concentrations measured by different platforms.

4.5.1 Luminex Bead Events

Bead events are particular to the Luminex platform and an adequate number of bead events are crucial for analysis. Higher bead counts can improve assay precision, whereas insufficient bead counts can render assay measurements unusable [281, 282, 259]. Low bead counts can therefore effectively waste precious samples. The excellent bead events of the mLuminex monoplex seen in Chapter 3 translated to a 25-assay multiplex from a different manufacturer. The mLuminex system delivered consistently more "functional" and "ideal" samples, classified with respect to bead counts, than corresponding assays on the pLuminex system from the same manufacturer. The superiority of the magnetic system is probably greater than depicted since all SF sample wells achieved $\geq$100 bead events for each analyte, but an "ideal" sample is defined as $\geq$2 replicate wells with $\geq$50 bead events. The high efficiency of magnetic separation during wash steps most likely maximises bead retention and prevents the formation and persistence of bead aggregates and complexes. On average, 10% of SF samples analysed by the pLuminex system were not "functional", which represents an unacceptable wastage of 1 in every 10 samples.

4.5.2 Precision

All three platforms demonstrated good intra-assay precision for the analysis of SF samples. The median intra-assay CV for all analytes was acceptable irrespective of platform and there were no differences between platforms in the proportion of SF samples with satisfactory CVs. However, there was a trend for lower CVs on the magnetic compared to polystyrene Luminex system although this was rarely significant. The most likely explanation for this is the superior bead events achieved by the magnetic bead assays [281, 282]. High intra-assay precision has been demonstrated for both Luminex and MSD for cytokine measurements in other biological fluids, but the inter-assay precision is often not as satisfactory [305, 306, 307, 289, 291, 292]. Inter-assay precision can be influenced by numerous factors such as reagent stability, errors in reagent preparation; variations in pH of buffers, temperature, wash procedure; pipetting errors; instrument calibration/performance fluctuations, complexity of the biological matrix and storage stability of samples. There can also be significant lot-to-lot and inter-laboratory differences in multiplex cytokine analysis, although the MSD platform my be more resilient to such variability [307, 293, 308]. The sources of inter-laboratory variation have been classified as pre-analytical (i.e. sample collection and storage); analytical (i.e. sample and reagent preparation and handling, implementation of assay protocol, and equipment maintenance and calibration); post-analytical (i.e. standard curve fitting, data analysis and quality control); and inherent to the kit (i.e. lot-to-lot variability) [308]. One of the limitations of this work is that inter-assay precision was not evaluated in an attempt to preserve precious SF sample aliquots for further analysis in future chapters.

4.5.3 Limits of Detection and Quantification

Consistent with previous comparison studies, MSD assays had the lowest LOD and LOQs compared to Luminex assays [307, 289, 291, 292, 294]. The pLuminex assays generally had lower LODs than their mLuminex counterparts, although this was not usually significant. The LOQs were often equivalent between Luminex platforms or lower on the pLuminex system, but this was rarely significant. Although not formally

tested in this study, the MSD platform is reported to have most linear signal output and wider dynamic range [289, 290, 291]. Potential reasons for MSD's superiority include: stronger avidity of antibody pairs; favourable antibody kinetics due to specific target epitopes; high sensitivity and lower background of electrochemiluminescence detection that also avoids time-dependent signal decay [309, 310, 292].

The LOD is the minimum concentration that can be reliably distinguished from blank measurement and can therefore be considered as an indicator of assay noise. It is often incorrectly used to describe analytical sensitivity, which is the ability to discriminate between 2 values and is given by the slope of the calibration curve [311]. The results suggest that the MSD platform has lower analytical noise than the Luminex platform. The LOQ is arguably a more useful parameter than LOD, since it represents the lower bound of the dynamic range for accurate and precise measurements. However, the LOQ value depends on the specific calibrator concentrations used for the standard curve. Therefore, LOQs should be compared with caution unless the standard curves have identical calibrator points. For example, 8-point standard curves for MSD assays were generated by 4-fold serial dilutions of a combined 2500pg/ml calibrator mix. For the Luminex platform, 8-point standard curves were produced by 3-fold serial dilutions of the stock calibrator, which contained different concentrations for each analyte. The lowest calibrator point of each respective standard curve defined the LOQs for IL-6, IL-8 and TNF-α on all platforms, but the absolute LOQ values (0.61pg/ml) were significantly lower for the MSD platform. A direct comparison of LOQs between Luminex systems can be made as they both shared identical calibration curves. There were fewer differences in LOQ (significant or non-significant) between Luminex platforms than LOD.

Disparities in LOD/LOQ concentrations and proportion of samples above these limits between the Luminex systems are difficult to explain. The 25-plex assay kits were from the same manufacturer and are assumed to be identical barring the nature of the microsphere beads, and thus 96-well plate and wash protocol. It is unlikely that the bead material is responsible for any observed differences. Potential explanations include subtle differences in bead-coupled capture and/or detection antibodies, reagents and diluents, and stock calibrator quality. The observations may also be (statistically) spurious since pLuminex limits were derived from only 3 serial plate-runs whereas

those for mLuminex were from 8 plate-runs. Fortunately, many differences in limits between bead systems were redundant as abundant or scarce markers were detectable and quantifiable in all or none of the SF samples, respectively. However, there were 2 types of anomalous observation: 1) the limits were equivalent, but the proportion of samples above limit was different; 2) the limits were different, and the proportion of samples above limit was opposite to expected. These may be due to differences between native marker epitopes within the sample matrix and lyophilised calibrator used for the standard curves generation, or the presence of binding factors in the clinical sample that interfere with epitope capture. In both anomalous scenarios there were more samples above limit on the mLuminex platform. It is possible that the greater bead retention with less aggregation and complex formation improved detection and quantification.

A common limitation to many platform comparison studies, including this work, is that LOD and LOQ is determined from measurements in assay diluent and not the biological matrix under investigation or the disease context of interest. There may be disparities in many aspects of assay performance when the same assay is conducted in different matrices [312]. Such parameters therefore represent the best possible performance of an assay, which are possibly worse in real world samples. The comparison of these artificial limits may not be useful to inform platform choice for a particular sample matrix or disease context. However, there was no single large volume OA sample for preparation of multiple standard curves and I wanted to avoid pooling of precious samples. Furthermore, the SF used for blanks and standard curves would ideally have to be depleted of all molecules under investigation, which is not feasible. The spike-recovery experiments would instead provide an estimate of matrix effects for the MSD and (m)Luminex platform. More important than absolute values of LOD and LOQ is the ability of an assay to detect and quantify markers in real samples in the context of disease. For example, the LOD and LOQs for IL-1β, IL-6, IL-8 and TNF-α are lowest for the MSD platform, but the proportion of SF samples >LOD and >LOQ for IL-6 and IL-8 was the same for all platforms. So all other parameters being equal, it would not matter which platform was used to detect and quantify IL-6 and IL-8 in OA SF. However, TNF-α and IL-1β were detected and quantifiable in significantly more SF samples using the MSD assays, which would be the platform of choice.

4.5.4 Spike Recovery and Accuracy

Both MSD and mLuminex platforms mostly demonstrated good accuracy as evidenced by satisfactory spike recoveries, with many analytes falling just short of acceptable, particularly at the higher spike concentration. For the 4 inflammatory markers in common, the MSD platform displayed overall better spike-recovery and therefore accuracy. The acceptable recoveries also suggest that there is little matrix interference during the analysis of most analytes on either platform. Though it must be stressed that the matrix in question is HAse treated SF and platform differences in spike recovery in neat SF samples were not tested. The findings in Chapter 3 suggest that the bead-based Luminex platform occurring in the fluid phase may be more liable to matrix interference than MSD due to aggregation and complex formation.

There were, however, suboptimal spike recoveries (accuracy) particularly at the lower concentration spikes on the Luminex platform. This may be due to the presence of binding factors, epitope denaturation or degradation in the SF sample preventing recognition. These effects may become saturated and less significant at the higher spike concentrations. Spike recovery for MIG and MCP-1 on the mLuminex platform produced some unexpected results. MIG was below detection in the unspiked sample, but recovery was >160% at both spike concentrations. This false positive signal may be due to cross-reactivity of the capture and/or detection antibodies, although the manufacturer's technical data sheet reports <1% cross-reactivity for all antibodies in the 25-plex. However, it also states that MIG recovery of >130% is possible in citrate plasma samples. MCP-1 had an acceptable recovery at the high spike concentration, but a negative recovery at the lower spike concentrations. This is most likely due to the large unspiked sample concentration, which was more than 20 times greater than the lower spike concentration. The final observed concentration was about 4% less than the unspiked measurement and may represent effective dilution during the spiking process. The CV between the expected and observed spike measurements is ~6%, which still represents good intra-assay precision.

One must take care not to over interpret and extrapolate the spike recovery experiments as they were conducted in only one patient's SF sample. In addition, recoveries were

tested at only 2 arbitrary spike concentrations that were different between platforms and for each analyte on the mLuminex system. The physiological range of the cytokines in healthy and OA states may not have been incorporated. Nonetheless, it is reassuring that the differences in recovery at higher and lower spike concentrations on either platform were not large since reproducible and consistent recovery is more important than achieving 100% recovery [263, 313]. Spike recovery may not be as good a measure of assay accuracy as perceived since manufacturer's calibrators seem to be optimised for their own respective assays, despite NICB quality certification. Chowdhury et al. showed that recoveries for MSD and Luminex assays in serum samples were excellent when spiked with the assays own calibrator, but generally poor and variable when spiked using calibrator from the alternate platform i.e. MSD calibrator used to spike sample measured by a Luminex assay and vice versa [289]. This highlights the need for universal biomarker reference calibrators to assess true assay accuracy.

4.5.5 ELISA Validation and Cross-platform Agreement

Agreement between two analytical methods allows their measurements to be used interchangeably, but does not suggest that either method has measured the true quantity or that either method has complete repeatability [314, 301]. ELISA is the most established immunoassay technique and is therefore often considered the gold standard for validating other methods. However, this does not imply that ELISA measurements represent the true quantity and therefore concordance with ELISA does not reflect accuracy. Cytokine measurements by different ELISA kits give considerably different results irrespective of sample matrix, whether the manufacturer's own calibrators or a common calibrator is used to generate standard curves [315, 316, 317]. The lack of consensus is due to differences in antibody pairs and their properties, and potentially different colorimetric detection substrates.

Although a discourse on analytical agreement and the statistical methods used to assess it are beyond the scope of this chapter, a few important points need to be made. Exact agreement in measured values (absolute agreement) is realistically impossible even when the same method is conducted twice on the same sample due to a degree of inherent error and imprecision. Therefore, pairwise hypothesis testing to measure agreement

is flawed. Correlation coefficients are also inappropriate since they only assess the linear or monotonic relationship i.e. association between measurements. Acceptable methods used to assess agreement include intra-class correlation, concordance, model II regressions and Bland-Altman analysis. Each particular tool has its own advantages and disadvantages with proponents and opponents [318, 319, 300, 302, 320]. What is an acceptable level of disagreement or bias between analytical methods depends on the context of the comparison and its implications. This study utilised 4 different methods to assess the agreement of multiplex platform measurements with ELISA and each other.

Generally speaking there was limited absolute agreement of concentration values produced by the multiplex platforms with ELISA and each other. There was usually a proportional bias, but no systematic bias in measurements. The MSD platform demonstrated the greatest agreement with IL-6 ELISA concentrations measurements and there was relatively good agreement for IL-6 and IL-8 concentrations between the Luminex bead systems. There are a multitude of studies that include a comparison of Luminex or MSD measurements with ELISA for a range of analytes in a variety of biological samples [1]. Such comparisons are often made in a flawed attempt to validate the accuracy of the multiplex assay. There are also reports with comparisons between Luminex kits, and one study compared MSD multiplex to monoplex kits [321, 331, 306, 336, 327, 294, 339]. There are fewer studies directly comparing the Luminex and MSD platforms to each other, and none in SF [307, 289, 291, 292, 294, 293]. Many studies incorrectly used correlation to comment on absolute agreement, but some studies did use more appropriate methods. Genuinely good absolute agreement between analytical methods has rarely been demonstrated, with proportional bias being the most common finding. Absolute measurements varied even for comparisons within the same multiplex system when comparing kits from different vendors or different types of kit (i.e. standard vs. sensitive or multiplex vs. monoplex) from the same vendor. This resonates with the poor absolute agreement, for most cytokines, between the polystyrene and magnetic Luminex kits from the same manufacturer.

[1] References 321, 305, 306, 322, 323, 324, 325, 326, 327, 294, 328, 293, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339

The reason for poor absolute agreement between assay measurements is primarily due to the quality and properties of the specific antibody pairs used for the sandwich immunoassay. Important characteristics include avidity and kinetics, which may be affected by temperature and pH; and the ability to recognise the target epitope in the biological specimen where the protein may be subject to binding factors, degradation, denaturation, polymerisation and modifications such as glycosylation. Multiplexing and the specific targets of the multiplex panel will also determine the quality of antibody pairs, which need to be optimised to maximise specificity and reduce cross-reactivity [340, 341]. The antibody pairs of any particular commercially available assay are therefore optimised for that specific kit, including reagents and calibrators. The importance of calibrators for standard curve generation must not be underestimated, since these too are optimised for the assay kit (antibody pairs and reagents) despite World Health Organisation/National Institute for Biological Standards and Control quality certification. The use of alternate calibrators instead of the kit's own can have significant effects on standard curves, spike-recovery and therefore the concentration measurements within the sample [315, 331, 289, 294]. There are presently no certified universal calibrator standards for biomarkers measurements. In theory, similar quantitative values are more likely if assays use identical antibody pairs, similar reagents and ideally a universal reference calibrator. The generally poor agreement between the p- and mLuminex assays, like their differences in LOD and LOQ, is difficult to reconcile as both were identical 25-plexes from the same vendor with assumed identical antibody pairs, reagents and calibrators. Subtle differences in these factors, as described previously, may offer some explanation.

The take-home message from this study and those preceding is that absolute agreement for concentration measurements between immunoassays is generally poor, but there is usually good relative agreement. Significant trends, patterns in the data and differences between cohorts are therefore preserved irrespective of the specific platform or assay. Fortunately, these represent the most clinically relevant and interesting findings of any study. However, quantitative comparisons between studies or the pooling of results from different studies is not possible unless identical techniques and methods have been used. Even so, there can be inter-laboratory (and thus inter-study) differences as discussed previously. This may pose a problem for biomarker research where biological

sample numbers are often small and collaborative work is essential.

At the time of writing, there were no head-to-head comparisons of Luminex and MSD performance for the analysis of SF from patients with knee OA. Parameters were examined using a within-subject experimental design to maximise power in the context of real SF samples. A standardised approach to standard curve fitting and interpolation was used as differences in proprietary software programs provided by platform manufacturers have a direct effect on results [258, 291]. Platform measurements were assessed for both absolute and relative agreement by 4 well-recognised methods. There were also several limitations to this work, which have been acknowledged during discussion of the study endpoints above. Furthermore, the effects of SF storage stability or freeze-thaw cycling on multiplex cytokine analysis was not assessed. Previous studies in other biological samples suggest cytokine stability of 1 year or more during storage at -80°C [342, 343, 344]. All patient SF samples were aliquoted after collection and processing to prevent the need for freeze-thaw cycling, which has been shown to variably effect cytokine measurements [342, 345, 346, 329, 347, 348, 349]. Ideally, positive and negative controls with known concentrations of markers of interest are required to confirm whether measurements in tested samples are real. However, obtaining such reference controls is often not possible i.e. difficulty in obtaining SF from a normal asymptomatic knee.

4.6 Conclusions

Null Hypotheses: Analysis of SF from patients with end-stage knee OA by the Luminex and MSD platforms demonstrate:

1. Acceptable assay performance in terms of limits of detection and quantification, intra-assay precision and spike recovery. - ***Accepted***
2. Equivalent assay performance between platforms in terms of limits of detection and quantification, intra-assay precision and spike recovery. ***- Rejected***
3. Absolute agreement of concentration measurements between platforms assessed by intraclass correlation, concordance, weighted Deming regression and Bland-Altman analysis. ***- Rejected***

This study goes some distance in validating the Luminex and MSD platforms for SF analysis with both platforms demonstrating acceptable precision and accuracy. However, there were some important differences between platforms. The MSD platform showed lower analytical noise (LOD) and better quantification limits. It has a similar protocol to the familiar ELISA and has a quick analyser read time. However, MSD multiplex assays are limited to a maximum of 10 markers [2]. The Luminex system allows higher multiplexing, but data is very dependent on bead events and the machine read time can thus often be long. The Luminex analyser requires regular careful maintenance and calibration. Through better bead retention, the magnetic Luminex system minimises potential wastage of precious SF samples and is to be recommended over its polystyrene counterpart.

Cross-platform comparisons of cytokine measurements suggest that biomarker measurements are at best only quasi-quantitative, with poor absolute agreement but excellent relative agreement. Once an analytical method has been used to measure a marker in a study, there needs to be strict consistency in methodology to facilitate evaluation and comparison of data. This should ideally encompass sample collection, storage

[2] Recent evolution of the MSD platform now offers multiplexing of up to 100 markers

and preparation; assay platform, manufacturer, specific kit, reagents and calibrator lot numbers. Platform selection for biomarker research will largely be determined by commercial availability, multiplexed panel, cost of assays, reagents and analyser, and ease of producing custom assays if required.

Part of this study was presented as a poster at the OARSI World Congress 2013, Philadelphia

Jayadev C, Snelling S, Price A, Hulley P. Multiplex Analysis of Osteoarthritic Synovial Fluid: a Comparison of Luminex and Mesoscale Discovery. Osteoarthr Cartil. 2013 Apr;21:S73–3.

Chapter 5

Anatomical Site of Synovial Fluid Sampling

5.1 Introduction

The burden of disease in osteoarthritis (OA) of the knee rarely has a uniform anatomical distribution [120, 163]. The most common pattern of knee OA is anteromedial knee OA (AMOA), which has a consistent topographical pattern of cartilage degeneration [250, 350]. Medial perimeniscal synovitis can be a common feature particularly when there is severe adjacent cartilage degeneration [83]. The lateral compartment of the knee has preserved articular cartilage with limited osteophytosis or synovitis. In comparison, tri-compartmental knee OA (TCOA) represents a global field change with the OA disease processes affecting all parts of the joint.

Synovial fluid (SF) is a non-Newtonian biological fluid with a complex non-homogenous matrix. Both flow and diffusion in such fluids are complicated and difficult to model. The synovial cavity of the knee joint is elaborate and convoluted due to the presence of recesses, bursae, pouches and intra-articular structures i.e. menisci and cruciates, [351, 352]. The flow of SF during knee motion is not simple and despite the free communication of SF between compartments of the knee joint, there is evidence for "physiological compartmentation" of the synovial space [352, 353, 354, 355]. This raises the potential for compartment-specific SF biology. The assumption of equilibrium in SF composition between all three knee joint compartments has never been formally

tested. Anatomical variations in SF analysis would have important practical implications for the level of consistency in SF sampling required for biomarker research; especially for unicompartmental OA.

Chapters 3 and 4 established the analytical requirements for reliable SF analysis. The importance of consistency in methodology, including sample collection was raised. This chapter explores whether the site of SF sampling is an important practical consideration. The objective of this study was to determine if the anatomical site of SF harvest affects synovial fluid biomarker measurements in patients with end-stage AMOA and TCOA.

5.2 Null Hypotheses

There is good agreement for SF concentration measurements of a range of soluble mediators between the 3 anatomical knee compartments in patients with:

1. Anteromedial knee OA (AMOA).
2. Tri-compartmental knee OA (TCOA).

5.3 Materials and Methods

5.3.1 Site-specific Synovial Fluid Collection

Synovial fluid (SF) was collected from the medial, lateral and patello-femoral (PF) knee compartments of 8 patients with end-stage AMOA and 8 patients with end-stage TCOA at the time of primary arthroplasty (total N=16). These samples were obtained from patients recruited specifically for this study. There were no restrictions placed on knee movements before surgery or during operative preparation and draping. SF collection was conducted after superficial dissection, but prior to arthrotomy to avoid contamination with blood. Approximately 250μl of SF was aspirated simultaneously and slowly from each joint compartment. The slow simultaneous collection of low volume samples was done to limit potential inter-compartmental SF flow. Samples were processed and stored as previously outlined.

5.3.2 Radiographic Assessment

Standardised pre-operative weight-bearing radiographs for the 16 patients were evaluated using the Kellgren & Lawrence and OARSI grading systems [114, 119]. Each of the three joint compartments was given a Kellgren & Lawrence grade (KLG), whereas the OARSI system was applied to the medial and lateral compartments only[1]. The KLG and the OARSI scores for joint space narrowing (JSN) and osteophytosis were compared between compartments for patients with AMOA (n=8) and TCOA (n=8). Median values (M) and interquartile ranges (IQR) are presented since the data are ordinal.

5.3.3 Synovial Fluid Analysis

All samples were analysed within 3 months of collection. SF samples were prepared for analysis and treated with hyaluronidase as previously described. Compartment-specific SF samples were analysed in triplicate for the following:

[1] As originally described [119]

- 8 pro-inflammatory cytokines (IL-1β, TNF-α, IL-6, IL-8, IL-15, IL-2, IL-12 & GM-CSF)
- 4 regulatory cytokines (IL-1RA, IL-4, IL-10 & IL-2R)
- 7 chemokines (Eotaxin, IP-10, MCP-1, MIG, MIP-1α, MIP-1β & RANTES)
- 3 isoforms of TGF-β (TGF-1β, TGF-β2 & TGF-β3)

Human cytokine 25-plex and TGF-β 3-plex magnetic-bead Luminex assays were conducted according to the manufacturer's protocol. The values for limits of detection (LOD) and quantification (LOQ) were as in Chapter 4 and Appendix C. Analytes with more than 75% measurements above LOQ were used for the repeatability analysis. The CV of technical replicates (intra-assay precision) was assessed with 20% used as the threshold of acceptability.

5.3.4 Internal Validation

SF samples from 8 patients with rheumatoid arthritis underwent identical preparation, processing and analysis as compartment-specific OA samples to provide internal validity of assay measurements. Heat maps were created using MeV Ver. 4.8.1 (TG4 Software). The inflammatory validation group were used as a positive control group against which other measurements were compared. Any potential differences between inflammatory and end-stage knee OA SF biology are explored in Chapter 9 and not discussed in this chapter.

5.3.5 Statistical Analysis

Radiographic Measures

Within-subject comparisons of KLG and OARSI scores were conducted separately for patients with AMOA and TCOA. The Friedman test with post-hoc Dunn's multiple comparison was used for KLG (3 compartments). The Wilcoxan signed rank test was used for OARSI scores (2 compartments).

Agreement of Compartment Measurements

The reproducibility of SF measurements between compartments was assessed for each analyte using the intraclass correlation coefficient (ICC) and coefficient of variation (CV).

Intraclass correlation coefficient: The absolute agreement between compartment-specific measurements was determined by the intraclass correlation coefficient (ICC) using a two-way mixed effects model for compartment measurements from individual patients i.e. $ICC_{3,1}$ [303]. The 3 compartments were considered the only "raters" of interest. The distribution of concentration measurements was assessed for normality by the Shapiro-Wilk test. If the distribution of measurements for an analyte was skewed in at least one compartment, data for that analyte in all 3 compartments were log-transformed and checked for a normal distribution prior to calculations of the ICC and its 95% confidence interval.

The ICC was determined separately for patients with AMOA (n=8) and TCOA (n=8), and interpreted using the following guide:

Value of ICC	Strength of agreement
<0.2	Poor
0.2 - 0.5	Fair
0.5 - 0.7	Moderate
0.7 - 0/8	String
>0.8	Almost Perfect

Negative ICCs are sometimes given by statistical software when the between-subject variation is relatively small compared to the within-subject variation. Negative ICC estimates suggest poor agreement, but cannot be reliably interpreted.

Cross-compartment coefficient of variation: The cross-compartment CV (cCV) of each analyte was calculated for each patient and expressed as a percentage.

$$cCV = \frac{\sigma_{comp}}{\bar{x}_{comp}}$$

where σ_{comp} is the standard deviation of measurements in the 3 knee compartments and $\bar{x}_{comp}$ is the mean. The median cCV for AMOA and TCOA patients was calculated for each analyte. The acceptable threshold for cCV was set at 20% because this is also the criterion for satisfactory intra- and inter-assay precision. The one-sample Wilcoxon signed rank test was used to compare the median cCV to the hypothetical (threshold) median of 20%. The Mann-Whitney test was used to compare the cCV between AMOA and TCOA patients for each analyte. Median values and non-parametric testing was used because cCV values have the potential to be 0, preventing log transformations in case of a non-normal distribution.

Reproducibility analysis, as described above, was conducted for analytes with quantifiable (>LOQ) matched compartment measurements in at least 6 (out of 8) patients. For analytes that did not meet this condition, the proportion of measurements >LOD and >LOQ were compared between compartments using Cochran's Q test for 3-matched proportions. $P<0.05$ was considered significant for statistical tests.

5.4 Results

5.4.1 Radiographic Assessment

The KL grade was not equal for medial, lateral and patello-femoral compartments in patients with AMOA ($p<0.001$) [Table 5.1]. The medial compartment had a significantly higher grade ($p<0.01$), with no significant difference between the lateral and PF compartments. The OARSI JSN and osteophyte scores were significantly greater in the medial than lateral compartment.

There was no significant difference in KL grade between compartments in patients with TCOA ($p=0.120$) [Table 5.1]. Likewise, the OARSI JSN and osteophyte scores were not significantly different between compartments.

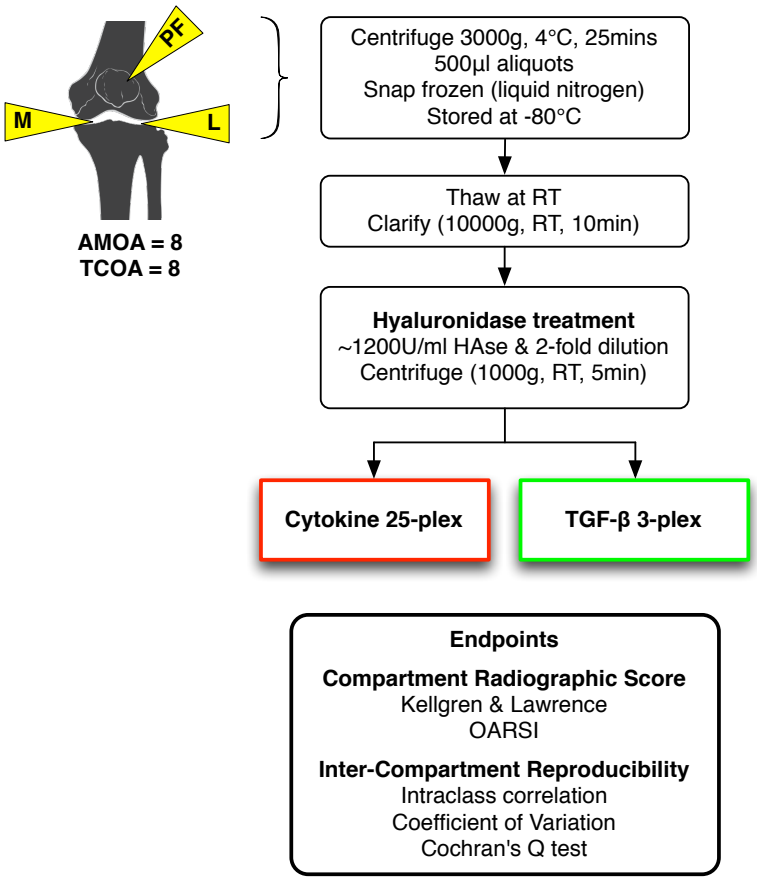


Figure 5.1 – Experimental design for assessing inter-compartment reproducibility of SF Analysis. SF was collected from the medial (M), lateral (L) and patello-femoral (PF) compartments of patients with AMOA (n=8) and TCOA (n=8). Samples underwent identical processing. SF was analysed for 9 pro-inflammatory cytokines (IL-1β, TNF-α, IL-6, IL-8, IL-15, IL-2, IL-12, IL-17 and GMCSF), 4 regulatory cytokines (IL-1RA, IL-4, IL-10 and IL-2R), 7 chemokines (Eotaxin, IP-10, MCP-1, MIG, MIP-1α, MIP-1β and RANTES) and TGFβ1, -β2 and -β3. Within-subject comparisons of KLG and OARSI scores between compartments were conducted for each patient group. The reproducibility of site-specific SF measurements was assessed for each analyte using the coefficient of variation (CV) and intraclass correlation coefficient (ICC). The acceptable threshold for cross-compartment CV (cCV) was set at 20%. An ICC >0.80 was considered excellent absolute agreement. Cochran's Q test for 3-matched proportions was used to compare number of measurements >LOD and >LOQ between compartments if the analyte had fewer than 6/8 quantifiable matched compartment-specific measurements.

AMOA n=8							
	Medial		Lateral		Patello-Femoral		p-value
KL Grade	**3**	(3 – 4)	**1**	(0 – 1.75)	**1.5**	(1 – 2.75)	<0.001
OARSI JSN	**3**	(2 – 3)	**0**	(0 – 0)			0.012
OARSI OP	**4**	(1.5 – 4)	**1**	(1 – 2)			0.030

TCOA n=8							
	Medial		Lateral		Patello-Femoral		p-value
KL Grade	**3**	(3 – 3.75)	**2.5**	(1 – 3)	**2**	(2 – 3)	0.120
OARSI JSN	**3**	(2 – 3)	**1.5**	(1 – 2.75)			0.182
OARSI OP	**4.5**	(3.25 – 5.75)	**3.5**	(1.25 – 5.5)			0.143

Table 5.1 – Comparison of radiographic assessments between compartments. The compartment Kellgren & Lawrence (KL) grade, OARSI joint space narrowing (JSN) and OARSI osteophyte (OP) scores were compared in patients with AMOA and TCOA. The medians and interquartile range (in brackets) are presented. In accordance with the original description there are no OARSI scores for the patello-femoral compartment. There were significant compartment differences in radiographic severity in AMOA patients but not TCOA patients.

5.4.2 Quality Control of SF Assays

In both AMOA and TCOA patient groups, there were 10 analytes above LOQ in compartment-matched samples for all patients: 4 pro-inflammatory cytokines (IL-6, IL-8, IL-12 and IL-15); 3 chemokines (Eotaxin, MCP-1 and IP-10); and the 3 isoforms of TGF-β. The remaining 12 analytes were quantifiable in all matched samples in fewer than 6 patients.

The median replicate CV (intra-assay precision) was between 3.43% and 12.39%. For 5 analytes (IL-6, IL-12, IP-10, IL-8, TGF-β2 and TGF-β3,) all samples had a replicate CV <20%. For the remaining 5 analytes, the number of samples with a replicate CV <20% range from 81% (IL-15) to 98% (TGF-β1).

5.4.3 Internal Validation

For the 10 analytes above LOQ in compartment-matched samples for all OA patients, the measured concentration was highest in inflammatory samples [Figure 5.2A]. This was significant for all analytes except Eotaxin and MCP-1. Except for IL-1β and IL-2, the remaining analytes were quantifiable in a significantly greater proportion of inflammatory samples [Figure 5.2B]. The differences confirm the validity assay measurements.

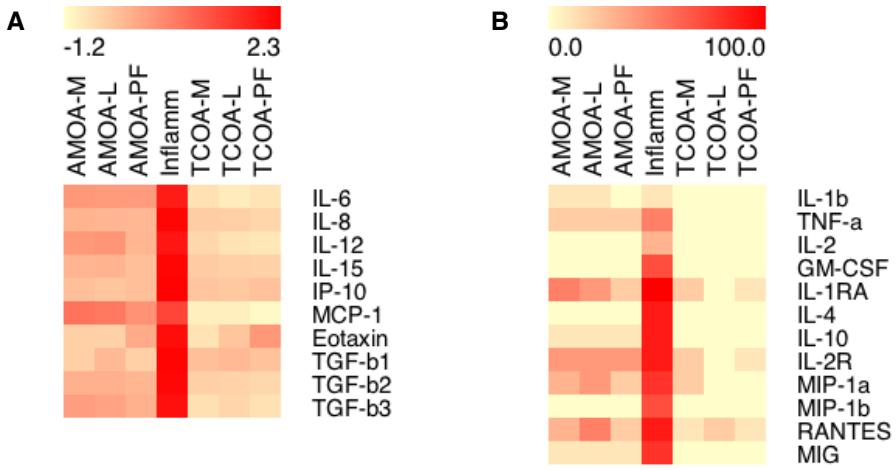


Figure 5.2 – Internal validation of compartment-specific SF analysis using rheumatoid arthritis SF samples (N=8). [A] The median concentration of 10 analytes that were >LOQ in all compartment-matched samples. Median analyte concentrations for each cohort were natural logarithm transformed and then standardised (mean centred and scaled to unit variance) **[B]** The proportion of quantifiable samples for analytes that were >LOQ in fewer than 6 compartment-matched samples.

5.4.4 Reproducibility of Compartment SF Analysis

Intraclass Correlation Coefficient

The ICC for compartment-specific measurements was >0.80 for all analytes in the AMOA group, even in the presence of a potential outlier, patient 7 [Table 5.2]. The lower bound of the 95%CI did not fall below 0.65 for any analyte. Seven (out of 10) analytes had ICCs >0.90. This suggests excellent absolute agreement between compartment measurements for analytes.

In the TCOA group, all analytes except IL-12 and TGF-β1 had ICCs >0.8 and 6 analytes had ICCs >0.90. The ICCs was 0.12 for IL-12 and 0.65 for TGF-β1. The lower bound of the 95%CI did not fall below 0.82 except for these two analytes and TGF-β3. An ICC of 0.12 and a negative lower bound of the ICC 95%CI for IL-12 suggest poor agreement.

The ICC for compartment measurements was re-calculated for the TCOA group without patient 12 identified previously [Table 5.3]. Exclusion of the outlier resulted in a large increase in ICC for both IL-12 and TGF-β1 to >0.80.

	AMOA		TCOA	
	ICC	**95%CI**	**ICC**	**95%CI**
IL-6	1.000	0.998 – 1.000	0.996	0.987 – 0.999
IL-8	0.985	0.949 – 0.997	0.990	0.936 – 0.998
IL-12	0.993	0.955 – 0.999	0.115	-0.172 – 0.608
IL-15	0.990	0.962 – 0.998	0.940	0.819 – 0.987
Eotaxin	0.910	0.740 – 0.979	0.955	0.864 – 0.990
MCP-1	0.989	0.964 – 0.998	0.975	0.919 – 0.994
IP-10	0.995	0.985 – 0.999	0.970	0.848 – 0.994
TGF-β1	0.876	0.657 – 0.971	0.653	0.261 – 0.907
TGF-β2	0.890	0.694 – 0.974	0.889	0.625 – 0.975
TGF-β3	0.879	0.661 – 0.972	0.811	0.527 – 0.954

Table 5.2 – Intraclass correlation coefficients (ICC) for absolute agreement between compartment-specific analyte measurements. A two-way mixed effects model was used for individual measurements (ICC3,1). The ICC and 95%CI are presented for each analyte in patients with AMOA and TCOA.

	Outlier Included [from Table 5.2]		Outlier Removed	
	ICC	**95%CI**	**ICC**	**95%CI**
IL-6	0.996	0.987 – 0.999	0.966	0.986 – 0.999
IL-8	0.990	0.936 – 0.998	0.990	0.940 – 0.998
IL-12	0.115	-0.172 – 0.608	0.801	0.400 – 0.959
IL-15	0.940	0.819 – 0.987	0.942	0.813 – 0.989
Eotaxin	0.955	0.864 – 0.990	0.957	0.860 – 0.992
MCP-1	0.975	0.919 – 0.994	0.978	0.922 – 0.996
IP-10	0.970	0.848 – 0.994	0.972	0.834 – 0.995
TGF-β1	0.653	0.261 – 0.907	0.833	0.540 – 0.965
TGF-β2	0.889	0.625 – 0.975	0.907	0.688 – 0.982
TGF-β3	0.811	0.527 – 0.954	0.844	0.655 – 0.977

Table 5.3 – Intraclass correlation coefficients (ICC) for absolute agreement between compartment-specific analyte measurements for TCOA patients with and without the potential outlier, patient 12. A two-way mixed effects model was used for individual measurements (ICC3,1). The ICC and 95%CI are presented for each analyte. There were large improvements in ICC for IL-12 and TGF-β1.

Cross-compartment Coefficient of Variation

The median cCV for all analytes was <15% in both AMOA and TCOA patient groups [Figure 5.3]. In the AMOA group, the cCV ranged from 3% (IL-6) to 11.6% (Eotaxin) [Figure 5.3A]. There were 5 analytes (IL-6, IL-15, MCP-1, IP-10 and TGF-β2) for which every patient had a cCV<20%. For the remaining 5 analytes (IL-8, IL-12, Eotaxin, TGF-β1 and TGF-β3) all but one patient had a cCV <20%. This outlier was the same patient (patient 7) for all of these markers except Eotaxin. Figure 5.4 shows that this patient had comparatively different lateral compartment measurements for these markers. The median cCV was significantly less than 20% for all analytes in AMOA patients, even when outliers were included.

In the TCOA group, the cCV ranged from 3.4% (IL-6) to 12.9% (Eotaxin) [Figure 5.3B]. There were also 5 analytes (IL-6, IL-8, MCP-1, IP-10 and TGF-β2) for which every patient had a cCV <20%. There were two patients for Eotaxin and one patient each for IL-12, IL-15, TGF-β1 and TGF-β3 that had a cCV above the satisfactory threshold. For IL-12, TGF-β1 and TGF-β3, this was the same patient (patient 12). Figure 5.5 shows the outlier patient 12 had a comparatively higher medial compartment concentration for these 3 analytes. In the TCOA patient group, the median cCV was significantly less than 20% for all analytes, except IL-12, IL-15 and TGF-β1. Exclusion of patient 12 as an outlier resulted in all cCVs being significantly less than 20%.

There was no significant difference in cCV between AMOA and TCOA patients for any analyte.

SF Measurements Below LOD and LOQ

For the remaining 12 analytes that had less than 6 (out of 8) quantifiable matched compartment measurements, the proportion of SF measurements >LOD or LOQ was not significantly different between compartments [Figure 5.6].

Four analytes (IL-2, GM-CSF, IL-4 and MIP-1β) were <LOQ in all compartment-specific samples in AMOA patients. In TCOA patients there were 8 analytes (IL-1β, TNF-α, IL-2, GM-CSF, IL-4, IL-10, MIP-1β and MIG) that were <LOQ in all compartment-specific

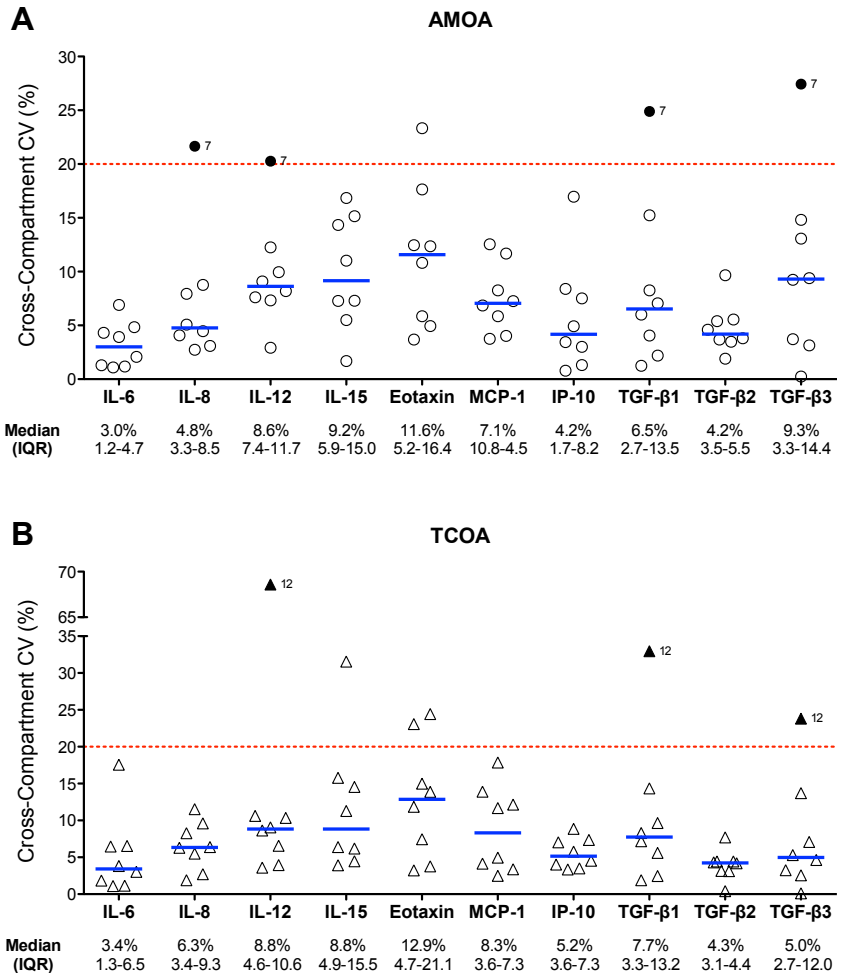


Figure 5.3 – Cross-compartment coefficient of variation (cCV) for SF measurements in [A] AMOA patients and [B] TCOA patients. A blue bar represents the median cCV for each analytes. The red dotted line at 20% represents the acceptable threshold for (inter- and intra-) assay precision. AMOA patient 7 (●) and TCOA patient 12 (▲) are highlighted as potential outliers. There was no significant difference in cCV between AMOA and TCOA patients for any analyte.

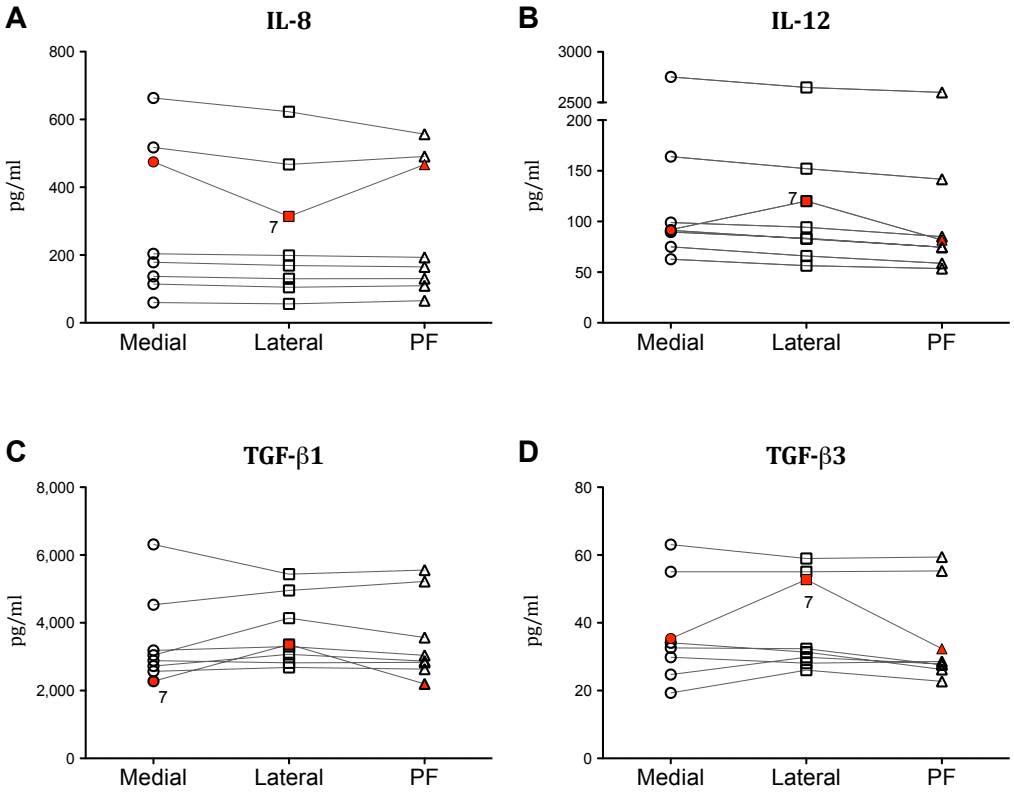


Figure 5.4 – Compartment-specific concentrations of [A] IL-8, [B] IL-12, [C] TGF-β1 and [D] TGF-β3 in patients with AMOA. Matched concentration measurements for the medial, lateral and PF compartment are shown. Note that Patient 7 had a comparatively different concentration for these 4 analytes in the lateral compartment (■).

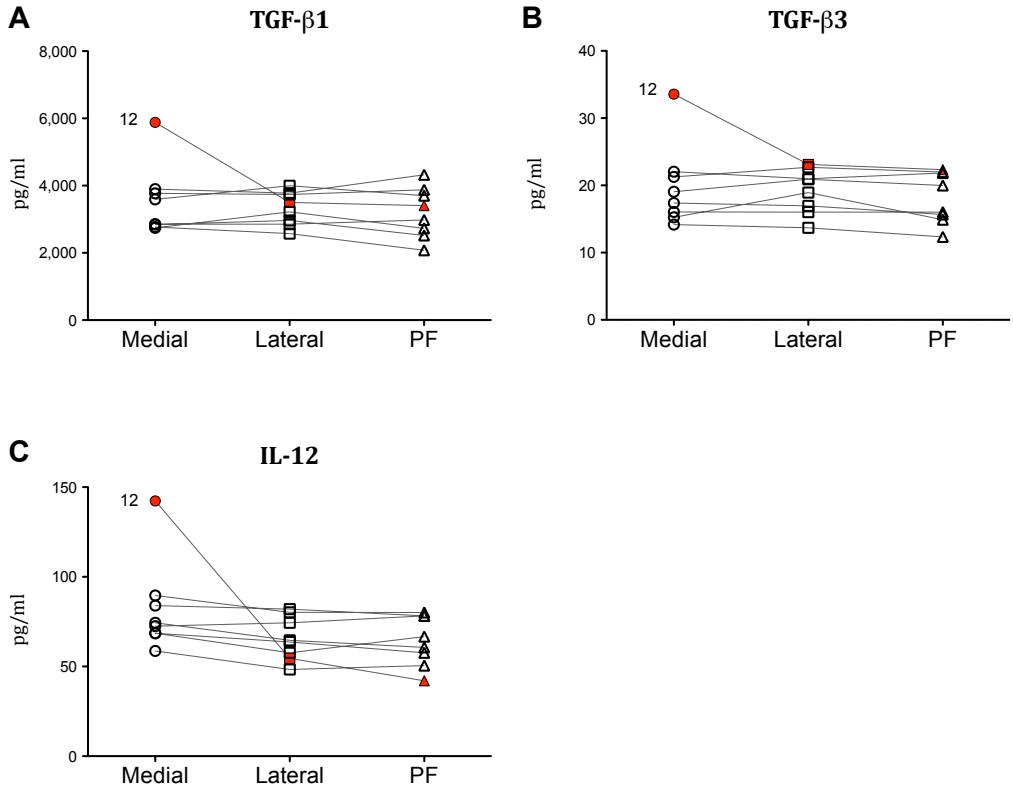


Figure 5.5 – Compartment-specific [A] TGF-β1, [B] TGF-β3 and [C] IL-12 concentrations in patients with TCOA. Matched concentration measurements for the medial, lateral and PF compartment are shown. Note that Patient 12 had a comparatively greater concentration of these 3 analytes in the medial compartment (●).

samples. Of these, 3 markers (IL-1β, TNF-α, IL-2 and IL-4) were undetectable in all samples.

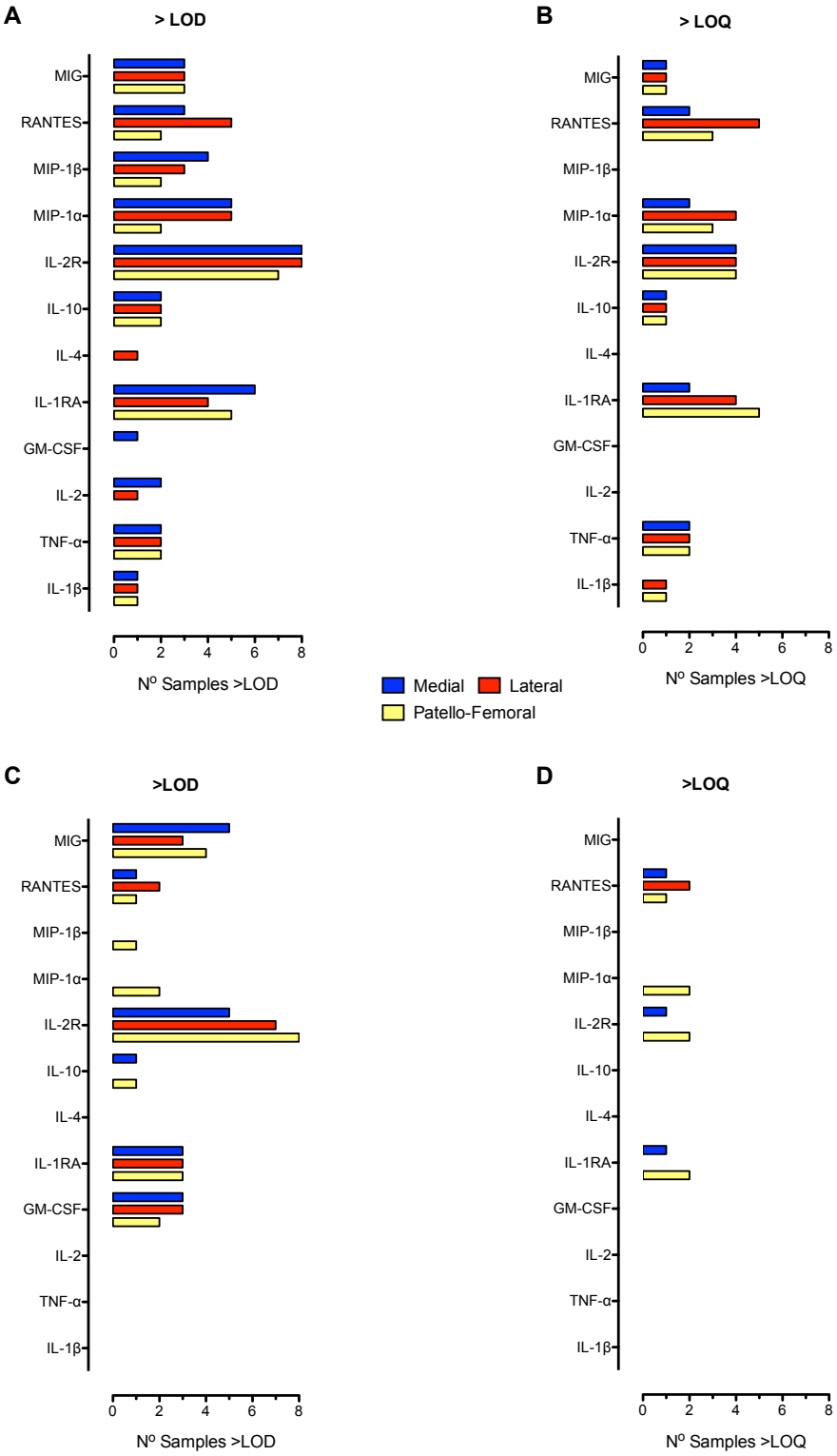


Figure 5.6 – Proportion of samples measurements above LOD and LOQ by compartment in [A & B] AMOA and [C & D] TCOA patients. The 12 analytes with fewer than 75% (6/8) matched measurements above LOQ are shown. Cochran's Q test was used to compare matched compartment proportions. There were no significant differences in the proportion of measurements >LOD or >LOQ between compartments for any analyte.

5.5 Discussion

The rationale for this study was to assess the agreement of SF analysis between the 3 anatomical compartments of the knee in the context of potential regional differences in OA disease pathology. The results demonstrate good reproducibility between compartment-specific SF measurements in both AMOA and TCOA. The internal validity of these measurements was confirmed using SF from patients with rheumatoid arthritis as positive controls. The anatomical site of sampling, therefore, does not affect SF analysis even when there are significant compartmental differences in structural burden of disease.

Pre-analytical factors can result in variability in SF biomarker measurements and include consistency of sample collection, handling and storage [308]. Conventional wisdom dictates that free communication within the synovial space coupled with the effects of joint movements result in efficient flow, mixing and diffusion of SF within different parts of the knee. However, there are both theoretical considerations and empirical observations that warranted this assumption being tested as it may have important practical implications for SF biomarker research.

Firstly, knee OA is a heterogeneous disorder with different, often overlapping, phenotypes based on clinical, radiographic, pathological and biological features [157, 158, 164, 356, 357]. The pattern of cartilage degeneration, osteophytosis and subchondral bone changes define the structural phenotypes of AMOA, lateral compartment OA, patello-femoral OA and TCOA [250, 358, 122]. There have been animal studies that suggest compartmentalisation of osteoarthritis pathophysiology [359, 360]. The severity of synovial activation in knee OA is variable and can also demonstrate differences in anatomical distribution [83, 361, 362, 77]. In particular, AMOA has been shown by arthroscopy and MRI to be associated with medial perimeniscal synovitis [83, 361]. An immunohistochemical study of compartment-specific synovial samples in patients with AMOA showed more severe synovitis and higher abundance of neuropeptides such as substance P, calcitonin gene-related peptide and neuropeptide Y in the medial synovium [363]. Substance P, in particular, has been implicated in the pathogenesis of OA [364]. Secondly, the knee joint is the largest and arguably

the most complicated synovial joint. Dye injection into the human knee and resin casts of rabbit knees in various positions of flexion-extension reveal a complex synovial cavity with thin intricate channels communicating between a large patello-femoral and smaller medial and lateral spaces [351, 352]. Furthermore, pressure studies in rabbit, cat and human knees suggest hydraulic compartmentalisation, at physiologic pressures, between anterior (patello-femoral) and posterior (medial and lateral) spaces, particularly at extreme ranges of motion [352, 353, 354, 355]. However, these studies were conducted in non-OA joints without a significant effusion. A recent arthroscopic cadavaric study into the dynamics of lavage showed that fluid flow to knee compartments is suboptimal and unevenly distributed irrespective of the configuration of portal positions [365]. The authors recommended entering each compartment individually for satisfactory lavage. Finally, the non-Newtonian nature of SF bestows complex rheological properties and diffusion characteristics that are beyond the scope of this thesis [366, 367, 11, 368, 369]. These factors taken together give rise to the theoretical notion of regional differences in SF biology.

Even with the assumption of perfect equilibrium of SF composition throughout the synovial space, one cannot expect identical site-specific measurements due to the constraints of intra-assay precision. As discussed in Chapter 4, correlations or repeated measure hypothesis tests are not appropriate methods to assess the reproducibility of compartment specific measurements. However, concordance coefficients, model II regression (i.e. weighted Deming regression) and Bland-Altman analysis used to test agreement in Chapter 4 were not used as these methods allow evaluation of only 2 sets of measurements. Consequently 3 pairwise comparisons for each analyte would be required to assess reproducibility across the joint compartments. The ICC is a common and valid method of assessing agreement between multiple “raters”, which in this study were the 3 anatomical compartments. The results suggested almost perfect absolute agreement for compartment-specific SF analysis. However, these high ICCs must be interpreted with restraint since wide variations in analyte levels between patients can lead to artificially high coefficients [320]. The cross-compartment CV allowed assessment of variability between site-specific measurements in terms analogous to intra- and inter-assay analytical precision. The variability was generally less than one may expect due to the limits of assay precision. In the absence of normal controls,

TCOA patients were used as positive controls where there was OA pathology in all 3 compartments compared to the predominantly unicompartmental disease in AMOA. However, TCOA patients still had the greatest radiographic severity in the medial compartment although this was not significant. This could explain the similar pattern of analyte cCV values for both patient groups [Figure 5.3]. Nonetheless, it was reassuring that there was still good reproducibility between compartments. Potential differences in marker levels between AMOA and TCOA were not explored as this assessed in Chapter 8. It remains possible that inter-compartmental flow and mixing occurred during site-specific SF collection despite best efforts to prevent it i.e. simultaneous, slow aspiration of small volumes. This may have masked real compartment-specific SF biology in both patient groups.

There are several limitations to this study that must be acknowledged. The number of patients in both patient groups was small. Although inflammatory samples were used to validate the authenticity of SF concentration measurements from OA patients, normal SF would have been more desirable. Radiographic scores were used as a surrogate for regional OA severity, and are weighted for (inferred) cartilage loss and osteophytosis. The extent and distribution of synovitis was not evaluated. Paradoxically, greater localised synovitis i.e. in the medial compartment of AMOA could result in greater localised SF reabsorption and clearance of marker levels [205]. Although a visual assessment of synovitis may have been possible during total knee replacement for TCOA, it would be difficult during routine minimal access surgery for unicompartmental knee replacement. Assessment of synovitis would therefore require intra-operative arthroscopy or pre-operative MRI. Likewise, no consideration was given to total SF volume or knee dimensions, which may influence flow and diffusion between compartments. Incorporation of such factors would require a much larger and complicated study design with additional radiological i.e. MRI, and biological i.e. SF/serum urea ratio assessments [220]. Finally, the anatomical reproducibility of the 22 analytes measured in this study has been extrapolated to all prospective SF markers. The evaluated cytokines, chemokines and growth factors have potential relevance to OA pathophysiology i.e. TGF-β in osteophytosis, but are by no means exhaustive. In particular, compartment-specific measurements of matrix enzymes and inhibitors were not tested. Ideally, potential biomarkers highlighted by exploratory studies should be

assessed for site-specific sampling stability as part of advanced pre-clinical validation.

5.6 Conclusions

Null Hypotheses: There is good agreement for SF concentration measurements of a range of soluble mediators between the 3 anatomical knee compartments in patients with:

1. Anteromedial knee OA (AMOA) **- Accepted**
2. Tri-compartmental knee OA (TCOA) **- Accepted**

There have been no previous studies exploring the potential for regional differences in SF biology within the same joint. The findings strongly suggest biological equilibrium in SF composition between joint compartments in end-stage knee OA. Since the different joint regions share a common synovial space this is perhaps unsurprising. Consequently, the anatomical site of SF collection is not an important practical implication for biomarker research in knee OA.

This study was presented as a poster at the OARSI World Congress 2013, Philadelphia

Jayadev C, Snelling S, Hulley P, Price A. The Anatomical Site of Synovial Fluid Sampling From the Osteoarthritic Knee: Does It Matter? Osteoarthr Cartil. 2013 Apr;21:S74–4.

Chapter 6

Short-term Temporal Stability of Synovial Fluid Analysis

6.1 Introduction

Short-term fluctuations in synovial fluid (SF) biomarkers may invalidate their utility for patient assessment, clinical decision-making and monitoring. Biomarkers should be dynamic and responsive to disease progression or treatment, but it is equally important they exhibit stability when there has been no progression or intervention [200, 199]. Ideally the intra-individual longitudinal variability of biomarkers should be relatively small, or at least well documented to ascertain minimally important clinical (or biological) changes.

The importance of short-term intra-individual variability is highlighted by the following hypothetical example. Protein-X is a burden of disease biomarker for knee OA and SF levels >60ng/ml are diagnostic of end-stage disease. It also has prognostic value in that SF levels >100ng/ml predict good long-term outcome from knee arthroplasty, whereas levels <100ng/ml are associated with higher rates of patient dissatisfaction. A patient with knee OA has a clinic appointment on 1st April. Clinical and radiographic assessments suggest severe OA and his SF Protein-X is 120ng/ml. Consequently, surgical management is discussed and the patient consents for knee arthroplasty. Due to an administrative error, the patient is given another clinic appointment on the 14th April. His symptoms, activity and lifestyle remain unchanged during this 2-week period.

Repeat clinical and radiographic evaluations were the same as previous. However, his SF Protein-X is 80ng/ml and consequently non-operative management is continued and a routine 6-month appointment is made. Two very different clinical decisions have been made due to intra-individual temporal variability in biomarker levels in the setting of clinical and structural constancy.

There are many studies investigating diurnal variations and effects of activity on marker levels, predominantly in serum and urine [370, 371, 372, 373, 374, 375, 376, 377]. There are very few studies with longitudinal SF analyses, especially in patients with OA [221, 222, 378]. This chapter aims to assess the short-term temporal stability of SF cytokine analysis in patients with end-stage knee OA who have no disease progression. In end-stage knee OA, burden of disease biomarkers should reflect the biological failure associated with the extent of clinical and structural (radiographic) severity. The SF biology should define a terminal profile and benchmark that measurements from earlier stages of disease progress towards. Consequently, one would expect the SF profile of end-stage knee OA to be stable over short periods of time, during which there is no further structural or clinical deterioration.

6.2 Null Hypothesis

There is low variability and good absolute agreement between SF analyses made over short periods of time in patients with end-stage knee OA.

6.3 Methodology

6.3.1 Ethics and Patient Recruitment

The longitudinal study was granted both regional (REC Ref: 11/SC/0549) and institutional ethical approval. All participants gave their written consent prior to their 1st SF collection and this was verbally reconfirmed on the day of surgery prior to further sample collection. Between June and September 2012, patients awaiting primary knee arthroplasty for primary knee OA were sent study information and approached in pre-operative assessment clinic (POAC). Criterial for exclusion from the study were as described in Chapter 2 [Section 2.4].

Forty patients were eligible, but only 12 patients volunteered to participate. The most common reason cited for non-participation was fear of postponement or cancellation of planned surgery and exacerbation of pain.

6.3.2 Sample Collection and Patient Assessments

Time-point 1: *Pre-operative assessment clinic*

Synovial fluid was collected from the index knee in the clinic room by a full aseptic technique. Approximately 500μl of SF was needle aspirated from the patello-femoral compartment with the needle entry point beneath the supero-medial pole of the patella. Sample collection was successful in 10 (out of 12) patient participants.

Time-point 2: *During surgery*

Approximately 500μl of SF was aspirated from the patello-femoral compartment after superficial dissection, but prior to arthrotomy to avoid contamination with blood. Samples were successfully collected from all 10 patients who donated fluid at time-point 1.

Samples were processed and stored as previously outlined [Section 2.6].

Each participant completed the Oxford Knee Score (OKS) and a Visual Analogue Pain score (VAS_{pain}) prior to sample collection at each time point. No restrictions were made on activity or lifestyle between the two time points.

Baseline standardised weight bearing radiographs of the index knee were taken at pre-operative assessment clinic and evaluated using the Kellgren & Lawrence and OARSI grading systems [114, 119]. However, it was neither clinically feasible nor ethically approved to repeat these on the day of surgery prior to operation. Since the macroscopic pathology of OA takes months or years to progress, it was assumed that there was no significant structural change in the short interval between time-points. The baseline scores are provided for descriptive purposes only.

6.3.3 Reference Group (External Validity)

The number of patients with clinical assessment and SF analysis at both time points (N=10) is small. To assess whether this small group was representative of patients undergoing primary knee replacement for end-stage knee OA, a further 15 patients were asked at random to complete an OKS questionnaire on the morning of their surgery. Their pre-operative OKS and standardised weight bearing radiographs of the index knee taken at POAC was retrospectively obtained. Radiographs were assessed using the Kellgren & Lawrence and OARSI grading systems. Patients at POAC do not routinely complete VAS_{pain}.

6.3.4 Synovial Fluid Analysis

All SF samples were analysed within 5 months of collection. SF samples were prepared for analysis and treated with hyaluronidase as previously described. SF samples were analysed in triplicate for the following:

- 8 pro-inflammatory cytokines (IL-1β, TNF-α, IL-6, IL-8, IL-15, IL-2, IL-12 and GM-CSF)
- 4 regulatory cytokines (IL-1RA, IL-4, IL-10 and IL-2R)

- 7 chemokines (Eotaxin, IP-10, MCP-1, MIG, MIP-1α, MIP-1β and RANTES)
- 3 isoforms of TGF-β (TGF-1β, TGF-β2 and TGF-β3)

Human cytokine 25-plex and TGF-β 3-plex magnetic-bead Luminex assays were conducted according to the manufacturer's protocol. All patient samples from both time-points were analysed on the same plate. The values for limits of detection (LOD) and quantification (LOQ) are given in Appendix C. Analytes with more than 70% (7/10) time-point matched measurements above LOQ were used for the repeatability analysis. The CV of technical replicates (intra-assay precision) was assessed with 20% as the threshold of acceptability.

6.3.5 Statistical Analysis

Clinical Assessments

The OKS was treated as a continuous variable, despite being the sum of 12 ordinal responses. The distribution of the OKS and VAS_{pain} at each time point was assessed for normality using the Shapiro-Wilk test and $\log_{10}$ transformed if necessary. Within-subject comparison of the OKS and VAS_{pain} between time-points was made using the paired two-tailed t-test.

The overall change in OKS (ΔOKS) and VAS_{pain} (ΔVAS) between time-points was calculated for each patient. The absolute change (magnitude, irrespective of direction) in OKS ($|\Delta OKS|$) and VAS_{pain} ($|\Delta VAS|$) was also calculated. These parameters were summarised by their median and interquartile range (IQR).

External Validity

Patient age, gender, knee OA phenotype, interval between POAC and surgery, baseline KL grade, baseline total OARSI score and the change in OKS (overall, ΔOKS, and absolute, $|\Delta OKS|$) were compared between the Experimental group and Reference group. The chi-square (Fisher's exact) test was used to compare proportions. Other

between-group comparisons were made using the two-tailed t-test (with Welch's correction for unequal variance) or Mann-Whitney test as appropriate.

Repeatability of SF Measurements

The repeatability of SF measurements between time points was assessed for each analyte using the coefficient of variation (CV), intraclass correlation coefficient (ICC), concordance correlation coefficient (CCC) and Bland-Altman analysis.

Interval coefficient of variation: The CV for a patient's SF measurements at each time point, the interval CV (iCV), was calculated for each analyte and expressed as a percentage.

$$iCV = \frac{\sigma_{interval}}{\bar{x}_{interval}}$$

where $\sigma_{interval}$ is the standard deviation of measurements taken at the 2 time-points and $\bar{x}_{interval}$ is the mean. The acceptable threshold for iCV was set at 20%, which is the threshold for acceptable intra- and inter-assay precision. The one sample Wilcoxon signed rank test was used to compare the median iCV to the hypothetical (threshold) median of 20%. Spearman rank correlation (one-tailed) was used to test for positive association between iCV and interval between SF measurements. Median values and non-parametric testing was used because iCV values have the potential to be 0, preventing log transformations in case of a non-normal distribution.

Absolute Agreement: The absolute agreement between compartment-specific measurements was determined by the ICC using a two-way random effects model for individual patient's time-point measurements i.e. $ICC_{2,1}$ [303]. The distribution of concentration measurements was assessed for normality by the Shapiro-Wilk test. If the distribution was skewed for an analyte for at least one time-point, data for that analyte at both time-points were $\log_{10}$ transformed and checked for normality prior to calculations of ICC or CCC and their 95% confidence intervals. ICCs were interpreted using the following guide:

Value of ICC	Strength of Agreement
<0.2	Poor
0.2 - 0.5	Fair
0.5 - 0.7	Moderate
0.7 - 0/8	Strong
>0.8	Almost Perfect

Negative ICCs produced by statistical software suggest poor agreement, but cannot be reliably interpreted.

The CCC was interpreted using the scale [304]:

Value of CCC	Strength of Agreement
<0.90	Poor
0.90 - 0.95	Moderate
0.95 - 0.99	Substantial
>0.99	Almost perfect

Bland-Altman analysis: Bland-Altman analysis included both classical difference vs. average plots and percentage difference vs. average plots. The latter was used to facilitate interpretation of average bias and 95% limits of agreement (LoA) between time-point measurements since the acceptability of these parameters, clinical or otherwise, is not known. The average of the time-point measurements is used as an estimate of the "true" concentration.

Repeatability analysis, as described above, was conducted for analytes with quantifiable (>LOQ) matched time-point measurements in at least 7 (out of 10) patients. For analytes that did not meet this condition, the proportion of measurements >LOD and >LOQ were compared between time-points using McNemar's exact test. $P<0.05$ was considered significant for statistical tests.

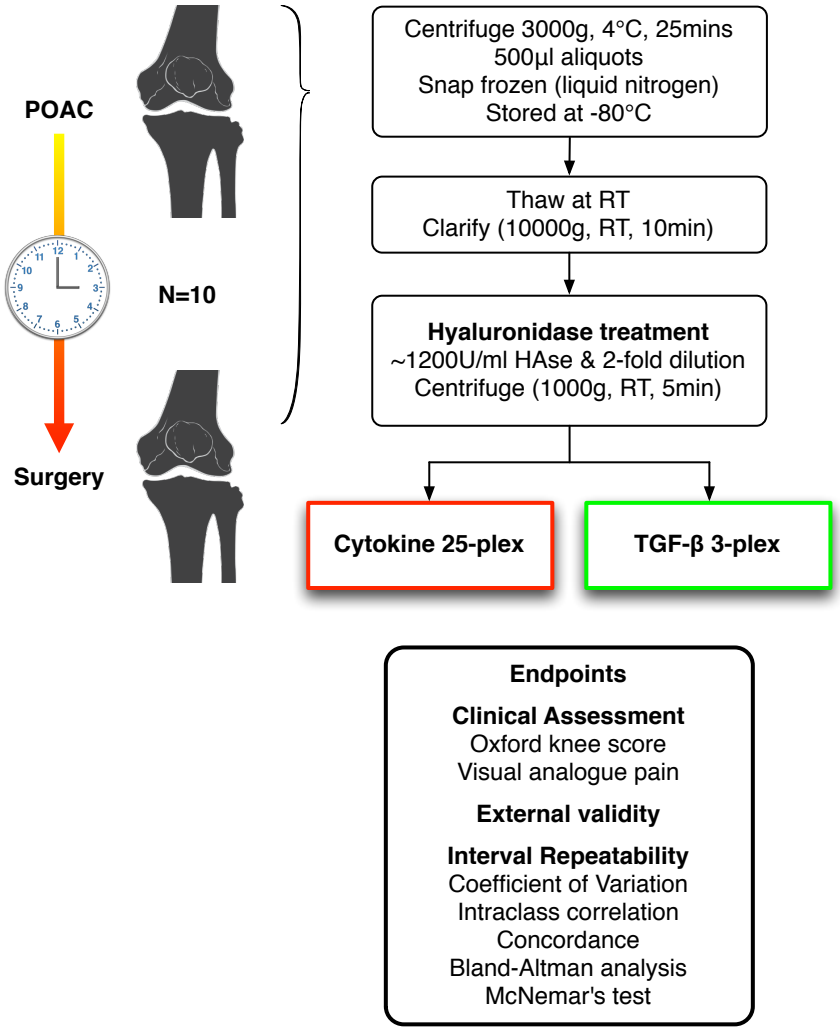


Figure 6.1 – Experimental design for assessing short-term temporal stability of SF analysis. SF was collected from patients with end-stage knee OA (N=10) at pre-operative assessment (POAC) and again at the time of primary arthroplasty. Samples underwent identical processing. Within-subject comparisons of Oxford Knee Score (OKS) and Visual Analogue Pain score (VAS_{pain}) between time-points were conducted for each patient. External validity was assessed by comparison of OKS (at POAC and surgery) and radiographic severity (Kellgren & Lawrence grade and OARSI score) to a Reference group (N=15) of randomly selected patients with end-stage knee OA. SF was analysed for 8 pro-inflammatory cytokines (IL-1β, TNF-α, IL-6, IL-8, IL-15, IL-2, IL-12 and GMCSF), 4 regulatory cytokines (IL-1RA, IL-4, IL-10 and IL-2R), 7 chemokines (Eotaxin, IP-10, MCP-1, MIG, MIP-1α, MIP-1β and RANTES) and TGF-β1, -β2 and -β3. The repeatability of matched SF measurements between POAC and Surgery was assessed for each analyte using the coefficient of variation (CV), intraclass correlation coefficient (ICC), concordance correlation coefficient and Bland-Altman analysis. The acceptable threshold for interval CV (iCV) was set at 20%. McNemar's test for matched proportions was used to compare number of measurements >LOD and >LOQ between time-points if the analyte had less than 7/10 quantifiable matched time-point measurements.

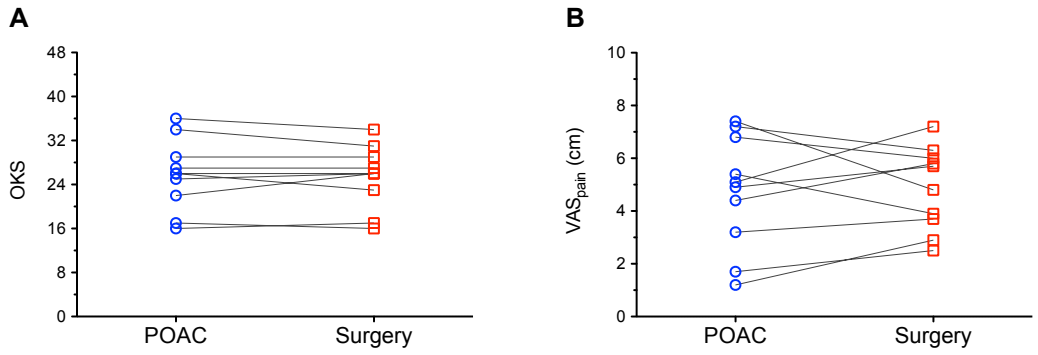


Figure 6.2 – Pairwise comparison of [A] Oxford Knee Score (OKS) and [B] Visual Analogue Pain Score (VAS_{pain}) at pre-operative assessment clinic (POAC) and on the morning of surgery. There was no significant difference in OKS (p=0.664) or VAS_{pain} (p=0.670, paired t-test).

6.4 Results

6.4.1 Clinical Assessments and External Validity

The median interval between POAC and surgery was 16.5 days (IQR 11.25 – 25). There was no significant difference in OKS or VAS_{pain} between time-points [Figure 6.2].

The median overall interval change in OKS was 0 (IQR -2.25 – 1) and the median absolute change was 1 (IQR 0 – 3) [Table 6.3]. The median overall interval change in VAS_{pain} was 0.7cm (IQR -1.1 – 1.6) and the median absolute change was 1.5cm (IQR 0.8 – 1.8).

Full clinical and radiographic assessments for the Experiment and Reference group are given in Table 6.1 and Table 6.2. There were no significant clinical or radiographic differences between the Experiment and Reference groups [Table 6.3].

6.4.2 Quality Control of SF Assays

Ten of the 22 analytes were above LOQ in both time-point samples for all patients: 4 pro-inflammatory cytokines (IL-6, IL-8, IL-12 and IL-15); 3 chemokines (Eotaxin, MCP-1 and IP-10); and the 3 isoforms of TGF-β [Figure 6.3 and Figure 6.4]. The remaining 12 analytes were quantifiable in paired samples for less than 7 patients.

Experimental Group									
Age	**Gender**	**Diagnosis**	**KLG**	**OARSI**	**OKS-1**	**VAS-1**	**Interval**	**OKS-2**	**VAS-2**
70	F	AMOA	3	9	25.00	3.2	13	26.00	3.7
77	M	AMOA	3	7	26.00	6.8	13	23.00	6.0
62	F	TCOA	4	19	26.00	7.4	25	26.00	4.8
52	F	AMOA	2	11	27.00	5.1	25	27.00	7.2
76	M	AMOA	3	9	34.00	1.2	27	31.00	2.9
70	M	TCOA	3	15	16.00	4.4	6	17.00	5.8
80	M	TCOA	3	13	22.00	5.4	22	26.00	3.9
71	M	AMOA	2	2	36.00	1.7	13	34.00	3.2
73	F	TCOA	4	18	17.00	7.2	6	16.00	6,3
65	M	AMOA	3	13	29.00	4.9	20	29.00	5.7

Table 6.1 – Clinical and radiographic assessments of Experimental group. Baseline radiographic Kellgren & Lawrence grade (KLG) and total OARSI scores were recorded at POAC. The Oxford Knee Score and Visual Analogue Pain Score at POAC (OKS-1 & VAS-1) and on the morning of surgery (OKS-2 & VAS-2) are presented. The time interval between POAC and surgery is presented in days. AMOA – anteromedial knee OA; TCOA – tri-compartmental knee OA.

Reference Group							
Age	**Gender**	**Diagnosis**	**KLG**	**OARSI**	**OKS-1**	**Interval**	**OKS-2**
80	F	AMOA	4	12	25	13	26
50	M	AMOA	2	8	23	20	24
81	M	AMOA	4	9	32	41	29
62	F	TCOA	4	14	16	6	14
67	M	TCOA	3	20	26	6	26
55	M	AMOA	4	12	13	9	14
58	F	AMOA	3	5	19	17	18
77	F	TCOA	4	14	10	12	12
75	M	AMOA	2	6	33	13	28
77	M	AMOA	2	4	18	34	21
72	M	AMOA	3	8	29	1	29
48	M	AMOA	2	5	18	5	17
64	F	AMOA	4	15	14	2	14
65	M	AMOA	3	12	29	3	28
72	F	TCOA	4	10	21	19	21

Table 6.2 – Clinical & radiographic assessments of Reference group. Baseline radiographic Kellgren & Lawrence grade (KLG) and total OARSI scores were recorded at POAC. The Oxford Knee Score at POAC (OKS-1) and on the morning of surgery (OKS-2) are presented. The time interval between POAC and surgery is presented in days. AMOA – anteromedial knee OA; TCOA – tri-compartmental knee OA.

	Experiment		Reference		P-value
Age	70	(64 – 75)	67	(61 – 73)	0.500
Phenotype	6 AMOA: 4 TCOA		12 AMOA: 4 TCOA		0.665
Gender	6 Male: 4 Female		9 Male: 6 Female		1.000
Interval	16.5	(11.25 – 25)	12	(5 – 19)	0.147
KLG	3	(2.75 – 3.25)	3	(2 – 4)	0.496
OARSI	12	(8.5 – 15.75)	10	(6 – 14)	0.556
OKS-1	25.8	(21.2 – 30.4)	21.7	(17.8 – 25.7)	0.154
OKS-2	25.5	21.5 – 29.5)	21.4	(18.0 – 24.8)	0.102
Δ OKS	0	(-2.25 – 1)	0	(-1 – 1)	0.888
|Δ OKS|	1	(0 – 3)	1	(0 – 2)	0.886
VAS-1	4.7	(3.2 – 6.3)			
VAS-2	4.9	(3.8 – 6.0)			
Δ VAS	0.7	(-1.1 – 1.6)			
|Δ VAS|	1.5	(0.8 – 1.8)			

Table 6.3 – Comparison of Experiment (N=10) and Reference patients (N=15) for external validity. There was no difference between Experiment and Reference patients with regards to demographics, interval between pre-operative assessment and surgery, radiographic severity, and clinical assessments. The mean and 95%CI are presented for Age, OKS and VAS_{pain}. The median and IQR are presented for interval, radiographic scores and change in (Δ)OKS and (Δ)VAS. KLG - Kellgren & Lawrence grade; OARSI - total OARSI scores; OKS - The Oxford Knee Score at POAC (OKS-1) and on the morning of surgery (OKS-2); ΔOKS – overall change in OKS between time-points; |ΔOKS| - absolute change in OKS between time-points; ΔVAS - overall change in VAS_{pain} between time-points; |ΔVAS| - absolute change in VAS_{pain} between time-points.

The internal validity of SF measurements by assays used in the study was previously confirmed using SF samples from rheumatoid arthritis as positive controls in Chapter 5.

The median replicate CV (intra-assay precision) was between 1.5% (IQR 0.6 – 3.4) and 11.9% (IQR 3.3 – 23.9). For 6 analytes (IL-6, IL-8, IL-12, TGF-β1, TGF-β2 and TGF-β3,) all samples had in intra-assay precision <20%. For the remaining 4 analytes, the number of samples with an intra-assay precision <20% range from 68% (IL-15) to 95% (Eotaxin, MCP-1 and IP-10).

6.4.3 Repeatability of Interval SF Measurements

Interval Coefficient of Variation

The median iCV was <15% for all 10 quantifiable analytes, ranging from 1.9% (TGF-β1) to 10.7% (IL-15) [Figure 6.5]. This was significantly less than 20% for all analytes ($p<0.05$, one sample Wilcoxon signed rank test). For 6 analytes (IL-12, Eotaxin, IP-10 and the three TGF-β isoforms), the iCV was <20% for every patient. There was one patient with an iCV >20% for IL-6, IL-8 and MCP-1 measurements, and two patients for IL-15. Patient 11 had an iCV >20% for IL-8, IL-15 and MCP-1.

There were no strong positive correlation between iCV and interval between SF measurements for any of the analytes (i.e. rho < 0.5 for all). None of the correlation coefficients were significant.

Intraclass Correlation and Concordance

Table 6.4 shows the ICCs and CCCs for absolute agreement between time-point SF measurements. The ICC was >0.80 for all analytes except TGF-β2 (ICC 0.75, 95%CI 0.30 – 0.93). For 8 of the 10 analytes (IL-6, IL-8, IL-12, IL-15, Eotaxin, MCP-1, IP-10 and TGF-β3), the ICC was >0.90 with the lower bound of the 95%CI >0.68. The CCCs reflected the ICCs, with the same 8 analytes achieving a CCC >0.9 and the lower bound of the 95%CI was >0.68. There was no significant difference in ICC or CCC between analytes.

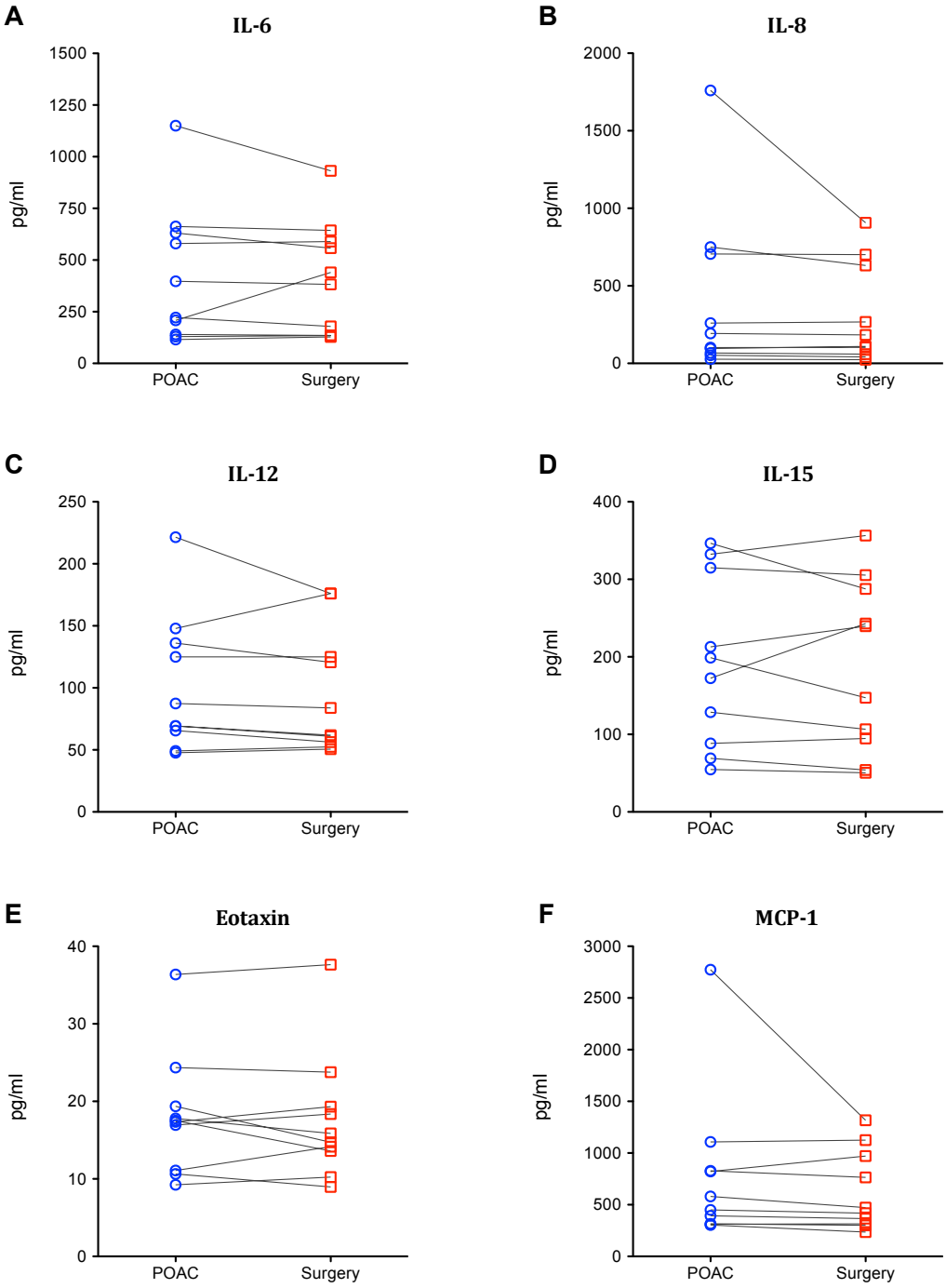


Figure 6.3 – **Intra-individual SF measurements** for **[A]** IL-6, **[B]** IL-8, **[C]** IL-12, **[D]** IL-15, **[E]** Eotaxin and **[F]** MCP-1 at preoperative assessment (POAC) and then at Surgery.

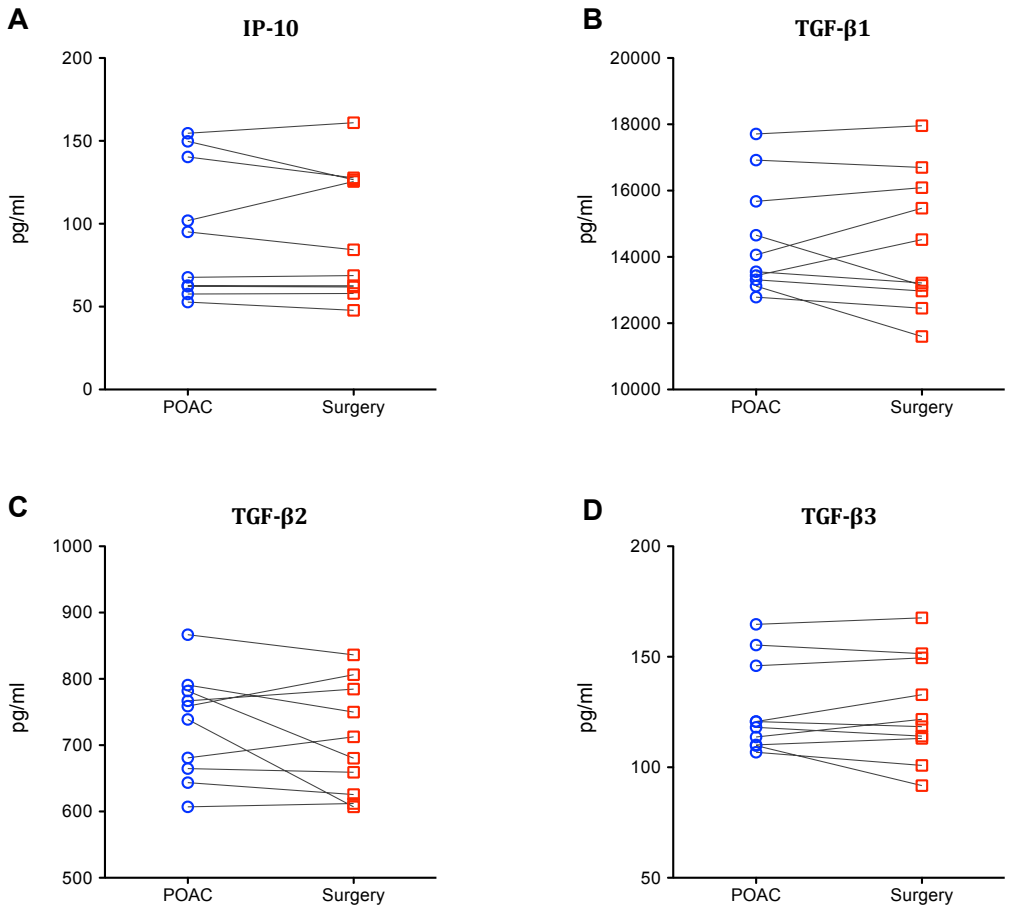


Figure 6.4 – Intra-individual SF measurements for **[A]** IP-10, **[B]** TGF-β1, **[C]** TGF-β2 and **[D]** TGF-β3 at preoperative assessment (POAC) and then at Surgery.

	ICC	95%CI	CCC	95%CI
IL-6	0.944	0.793 – 0.986	0.938	0.780 – 0.984
IL-8	0.983	0.925 – 0.996	0.981	0.932 – 0.995
IL-12	0.969	0.885 – 0.992	0.965	0.870 – 0.991
IL-15	0.962	0.860 – 0.990	0.958	0.849 – 0.989
Eotaxin	0.918	0.712 – 0.979	0.910	0.684 – 0.977
MCP-1	0.919	0.707 – 0.979	0.911	0.718 – 0.974
IP-10	0.970	0.888 – 0.992	0.967	0.873 – 0.992
TGF-β1	0.882	0.596 – 0.969	0.870	0.604 – 0.962
TGF-β2	0.747	0.299 – 0.929	0.726	0.250 – 0.920
TGF-β3	0.911	0.684 – 0.977	0.902	0.696 – 0.971

Table 6.4 – Intra-class correlation coefficients (ICC) and concordance correlation coefficients (CCC) for absolute agreement between analyte measurements taken at POAC and Surgery. A two-way random effects model was used for individual measurements ($ICC_{2,1}$). The coefficients and 95%CI are presented for each analyte.

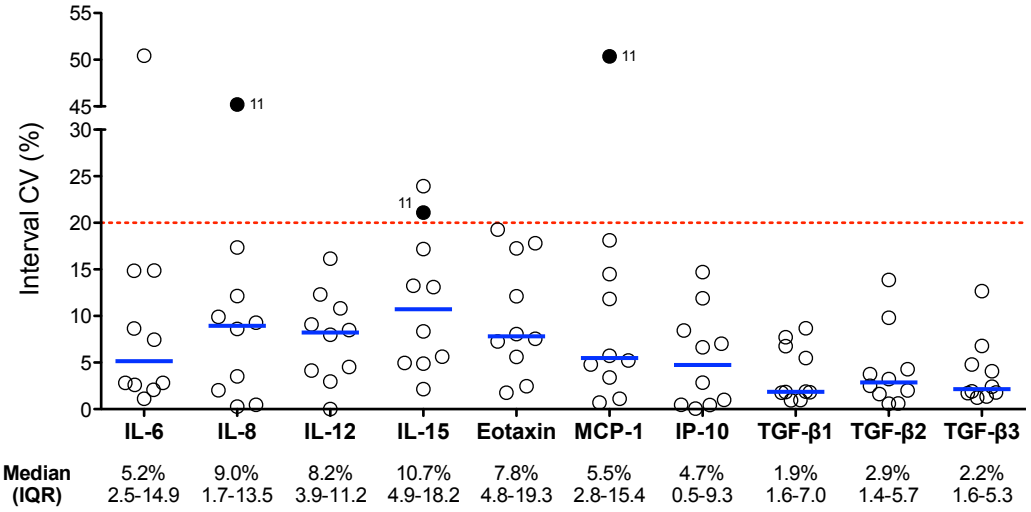


Figure 6.5 – Interval coefficient of variation (iCV) for SF measurements. A blue bar represents the median iCV for each analytes. The red dotted line at 20% represents the acceptable threshold for (inter- and intra-) assay precision. All analyte iCVs were significantly less than 20% ($p<0.05$, one sample Wilcoxon signed rank test). There was no significant difference between analyte iCVs (Kruskal-Wallis test). Patient 11 (●) is highlighted as a potential outlier.

Bland-Altman Analysis

The average bias and limits of agreement (LoA) between time-point measurements were variable across analytes [Figure 6.6 to Figure 6.9]. The mean bias ranges from 0.4pg/ml (Eotaxin) to 160pg/ml (MCP-1). The LoA were narrowest for Eotaxin (5.5pg/ml to -4.7pg/ml) and widest for TGF-β1 (1989pg/ml to 1767pg/ml). The significance of these concentrations is more easily interpreted in terms of percentage bias (%bias) between measurements. The mean %bias was <5% for all analytes except IL-8 and MCP-1. The %bias ranges from 0.8% (TGF-β3) to 12.5% (MCP-1). The corresponding LoA were narrowest for TGF-β1 (14.8% to -12.6%) and widest for IL-6 (49.3% to -52.9%). However, for all analytes the 95%CI for mean bias or %bias included zero (N=10).

The reproducibility of IL-8 and MCP-1 measurements was improved by exclusion of patient 11, who was identified as an outlier according to iCVs [Figure 6.5]. For IL-8: bias 14.9pg/ml with LoA from 92.8pg/ml to -62.9pg/ml; and %bias 6.4% with LoA from 29.2% to -16.5%. For MCP-1: bias 16.0pg/ml with LoA from 159.6pg/ml to -127.7pg/ml; and %bias 6.0% with LoA from 30.2% to -18.3%.

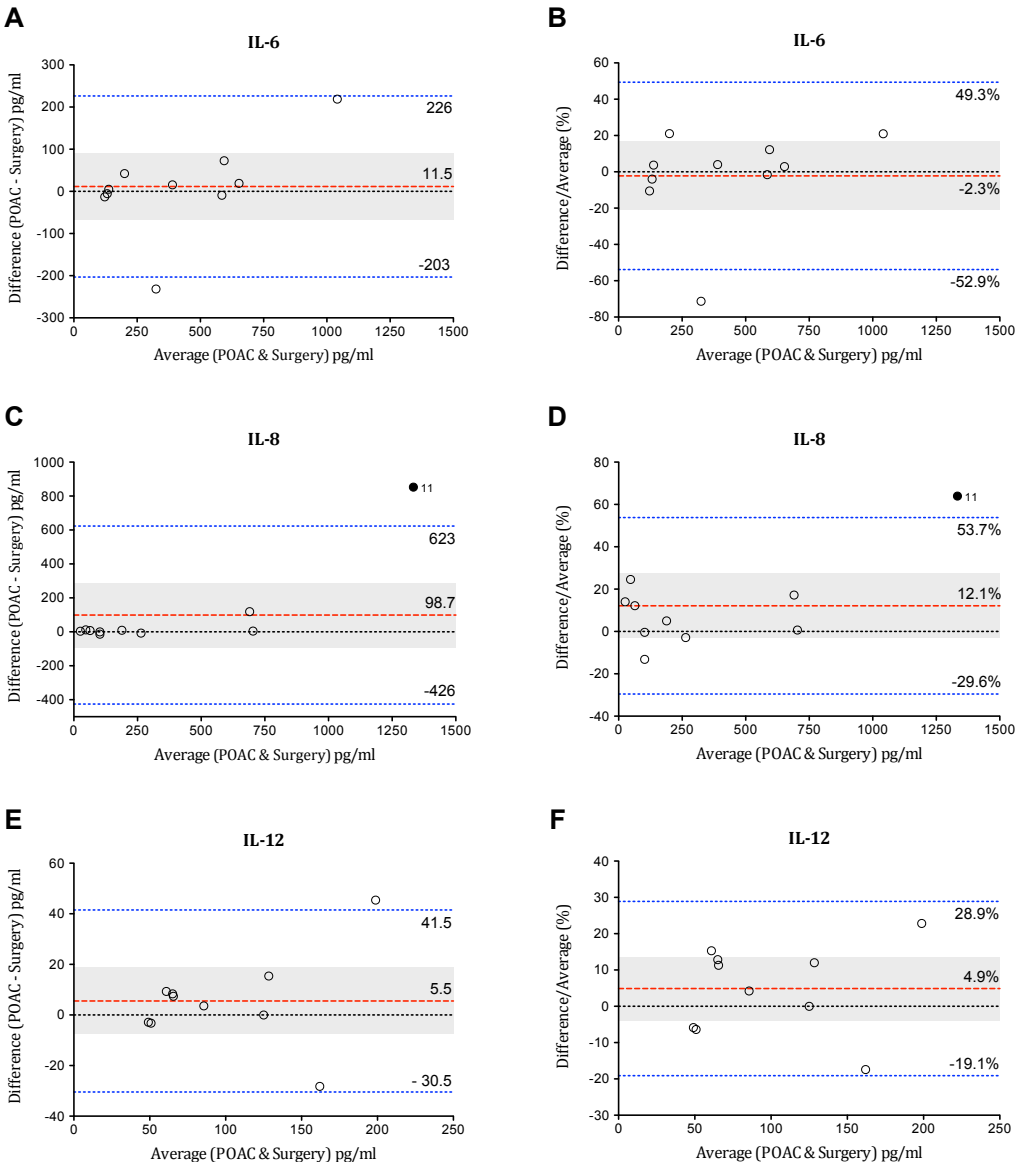


Figure 6.6 – Bland-Altman plots for agreement of [A & B] IL-6, [C & D] IL-8 and [E & F] IL-12 measurements taken at pre-operative assessment (POAC) and Surgery. The left-hand plots **[A, C & E]** are of difference vs. average measurements. The right-hand plots **[B, D & F]** are of percentage difference vs. average measurements. The black line represents zero absolute bias. The mean bias (red hashed line) with 95%CI confidence interval (grey shading), and 95% limits of agreement (blue hashed line) are shown. Patient 11 (●), who was identified as a potential outlier for IL-15 and MCP-1 according to iCV in Figure 6.5, is highlighted.

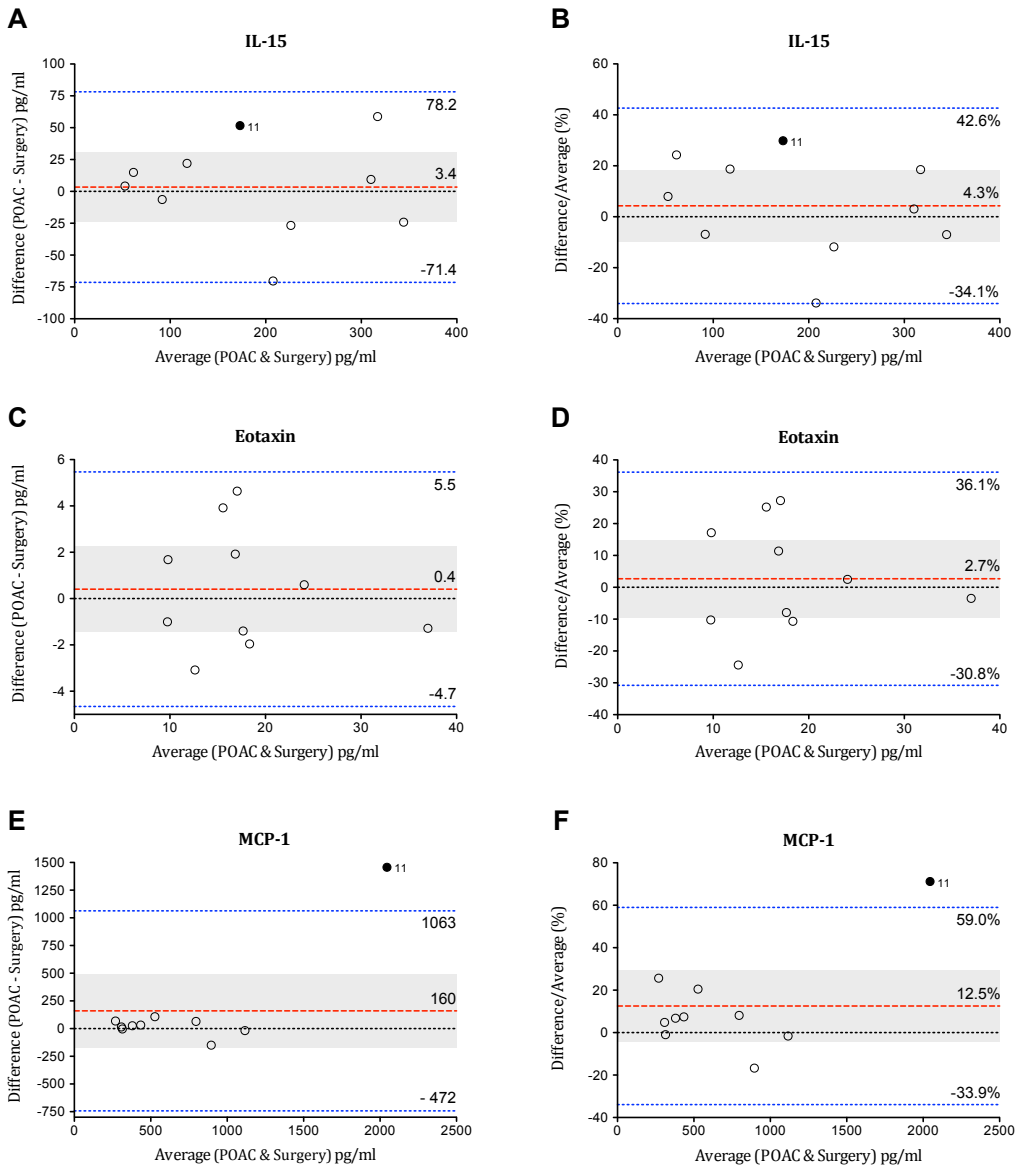


Figure 6.7 – Bland-Altman plots for agreement of [A & B] IL-15, [C & D] Eotaxin and [E & F] MCP-1 measurements taken at pre-operative assessment (POAC) and Surgery. The left-hand plots **[A, C & E]** are of difference vs. average measurements. The right-hand plots **[B, D & F]** are of percentage difference vs. average measurements. The black line represents zero absolute bias. The mean bias (red hashed line) with 95%CI confidence interval (grey shading), and 95% limits of agreement (blue hashed line) are shown. Patient 11 (●), who was identified as a potential outlier for IL-15 and MCP-1 according to iCV in Figure 6.5, is highlighted.

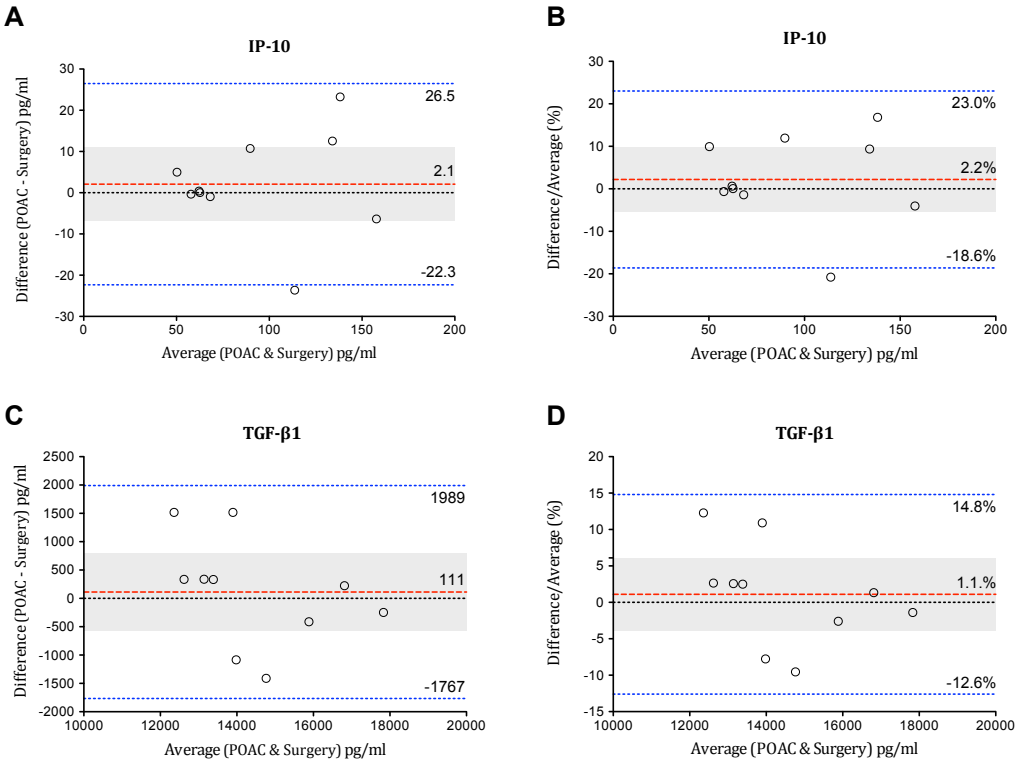


Figure 6.8 – Bland-Altman plots for agreement of [A & B] IP-10 and [C & D] TGF-β1 measurements taken at pre-operative assessment (POAC) and Surgery. The left-hand plots [A & C] are of difference vs. average measurements. The right-hand plots **[B & D]** are of percentage difference vs. average measurements. The mean bias (red hashed line) with 95%CI confidence interval (grey shading), and 95% limits of agreement (blue hashed line) are shown.

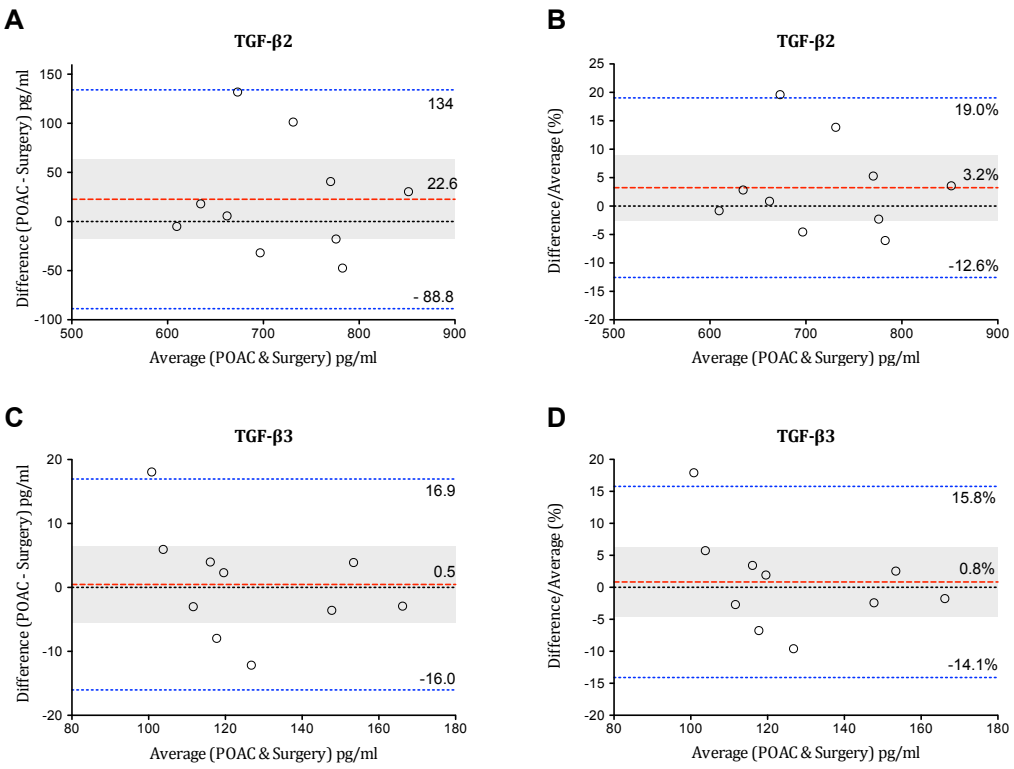


Figure 6.9 – Bland-Altman plots for agreement of [A & B] TGF-β2 and [C & D] TGF-β3 measurements taken at pre-operative assessment (POAC) and Surgery. The left-hand plots [A & C] are of difference vs. average measurements. The right-hand plots **[B & D]** are of percentage difference vs. average measurements. The black line represents zero absolute bias. The mean bias (red hashed line) with 95%CI confidence interval (grey shading), and 95% limits of agreement (blue hashed line) are shown.

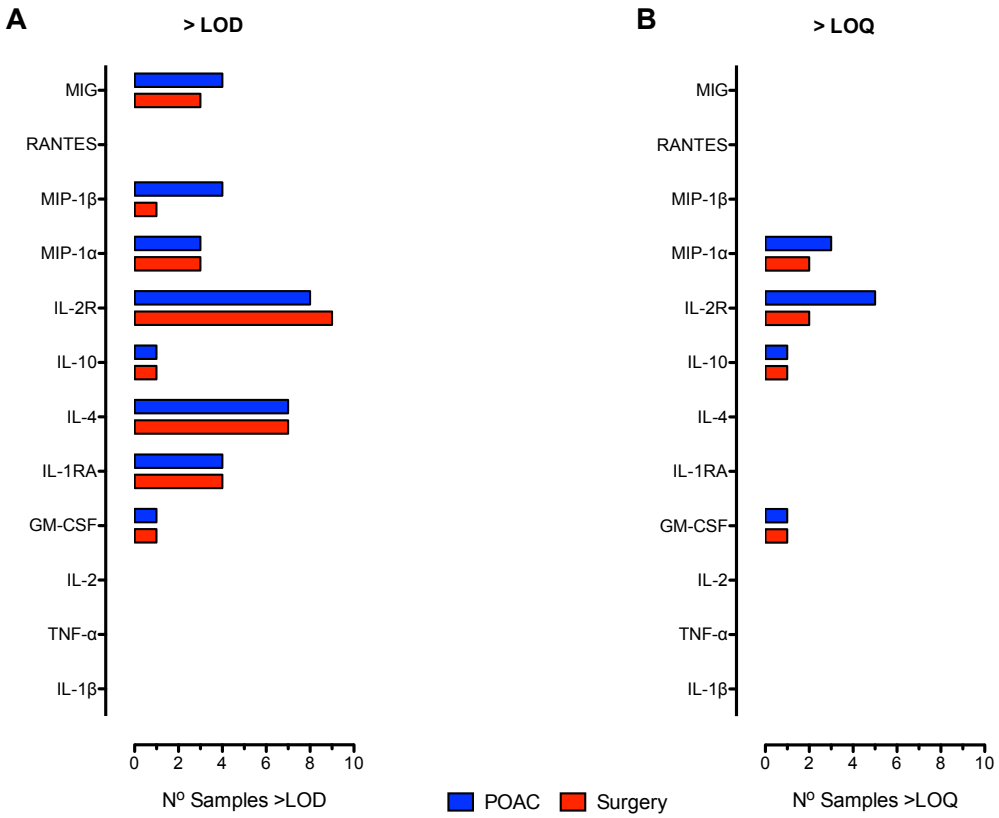


Figure 6.10 – Proportion of samples measurements above [A] LOD and [B] LOQ by time point. The 12 analytes with fewer than 70% (7/10) matched measurements above LOQ are shown. McNemar's exact test was used to compare matched proportions. There were no significant differences in the proportion of measurements >LOD or >LOQ between time-points for any analyte.

SF Measurements Below LOD and LOQ

For the remaining 12 analytes that had less than 7 (out of 10) quantifiable matched time-point measurements, the proportion of SF measurements >LOD or LOQ was not significantly different between POAC and Surgery [Figure 6.10].

Six analytes (IL-1β, TNF-α, IL-2, IL-RA, IL-4, MIP-1β, RANTES and MIG) were <LOQ in all samples at both time-points. Of these, 3 markers (IL-1β, TNF-α, IL-2 and RANTES) were undetectable in SF samples taken at either time point.

6.5 Discussion

Short-term fluctuations in biomarker measurements unrelated to disease progression or treatment would invalidate the utility of biomarkers for patient assessment. The aim of this study was to assess the temporal stability of SF cytokine levels in patients with end-stage knee OA between pre-operative assessment and surgery, during which time there was no clinical or structural progression. The results demonstrate acceptable intra-individual temporal variability and good agreement in a wide range of cytokine measurements taken approximately 2 weeks apart. However, there is most likely a degree of inter-individual variability between time-point measurements that varies between markers, but may not be clinically or biologically important.

The short-term stability of SF biology was assessed in the context of end-stage primary knee OA, which by virtue of its severity should not be associated with any significant progression, either clinical or radiological, over a matter of weeks. Therefore the burden of disease reflected by SF analysis should also remain constant. The absence of any significant differences in OKS or VAS_{pain} between time-points supports the lack of clinical progression. Unchanged radiographic severity was assumed based on the slow natural history of OA and is acknowledged as a limitation of the study. The stability of SF cytokines over the short time interval was tested by a number of statistical methods: CV, ICC, CCC and Bland-Altman analysis for quantifiable markers, and McNemar's test for markers below LOQ. The interval CV allowed assessment of variability between time-points in terms analogous to intra- and inter-assay analytical precision. It is reassuring that the variability was significantly less than one may expect due to the limits of assay precision. The absolute agreement between time-point analyses was strong to almost perfect according to the ICCs. However, ICCs can be falsely raised when there is substantial between-patient variability in analyte levels [320]. The concordance reflected the ICCs but was less favourable, being moderate to substantial, according to McBride's interpretation [304]. Although it is possible that this interpretation may be too stringent for biological data when both sets of measurements are susceptible to error and analytical imprecision, the CCC is less susceptible to between-patient heterogeneity than the ICC. Therefore, it may represent a degree of real variability in SF biology over short intervals even in end-stage OA. This notion is supported by the

Bland-Altman analysis. Each analyte did have an average bias and limits of agreements that varied by analyte in terms of magnitude and width, respectively. While the 95%CIs for the mean bias (or %bias) included zero for all analytes suggesting overall acceptable agreement between time-point measurements, this may be a consequence of the low sample numbers or the influence of outliers (i.e. IL-8 and MCP-1). Whether these biases and limits of agreement are clinically or biologically important is unknown. A fundamental principle in Part I of this thesis has been that perfect absolute agreement between measurements is usually impossible; whether between analytical methods, anatomical sites and now time points.

Identical SF biology at two separate time points is unlikely, even over a short interval in the relative steady state of end-stage knee OA and allowing for analytical imprecision. There are several possible reasons for variability in SF marker concentrations over the short interval between POAC and surgery. Sampling conducted at different times of day may be subject to diurnal variations in marker levels. This has been demonstrated for serum and urine analysis in patients with OA, but there are no specific studies for SF analysis [373, 374, 376]. The level of activity, in particular, and food intake prior to sampling are also reported to influence biomarker levels and may contribute to diurnal variations [370, 371, 372, 375, 377]. The study by Roos et al. is highlighted because it included both SF and serum markers of cartilage matrix metabolism before and after exercise. SF markers were relatively stable with no significant change after exercise, although there was a general rising trend. However, the study was conducted on healthy athletes and may not be applicable to patients with knee OA. Activity and food intake are particularly pertinent to sample collection at surgery since patients are fasted and have reduced activity. Additional pre-analytical causes of variability associated with surgical sample collection may include elevated circulating glucocorticoid due to anxiety and surgical stress response, and the use of a thigh tourniquet. It is also possible that recent needle aspiration at POAC could temporarily alter SF biology measured subsequently at surgery. Although a significantly more invasive procedure, arthroscopy has been shown to significantly reduce SF levels of the keratan sulphate, but not chondroitin sulphate, in patients with knee OA [379]. Despite the potential causes for temporal variability in SF biology given above, it is reassuring that the spread of bias (from Bland-Altman analysis) of all markers in this study spans either side of

zero indicating that the variability is not associated with any systematic increase (or decrease) in marker levels over the time interval.

At the time of writing, the present study is the first to focus on the short-term stability of potential SF biomarkers in knee OA. There is a notable paucity of studies examining longitudinal changes in SF biology in human subjects. The majority of OA biomarker investigations, including the Osteoarthritis Initiative concentrate on serum, plasma or urine. At the time of writing, there was only one previous study explicitly evaluating the longitudinal stability of SF markers in patients with knee OA [222]. Lohmander et al. measured aggrecan fragments and matrix enzymes (MMP-1, MMP-3, and TIMP-1) in SF of 52 patients with early knee OA sampled at 8 time-points over 1 year. The within-patient variability of the SF markers, given as the average CV over the 1-year period, ranged from 18% (aggrecan fragments measured by dye precipitation) to 71% (MMP-1 measured by ELISA). However, the majority of patients had a recent history of meniscal injury and/or cruciate ligament rupture and the authors did not provide data for SF variability over shorter time intervals. One may reasonably expect significant inter-individual changes in SF biology over 1 year in patients with early OA, especially in the context of these injuries.

The invasive nature of sample collection is a commonly encountered challenge to longitudinal studies involving SF biology. Consequently, most longitudinal biomarker studies in OA favour peripheral blood or urine analysis. Recruitment for the present study was difficult and only 30% of patients approached volunteered to participate. Needle aspiration in clinic and potential consequences, such as pain and delay or cancellation of surgery were the primary reasons for poor recruitment. However, this should not deter attempts to evaluate the longitudinal stability of SF biomarkers. The neuroscience literature can be used as motivation with several studies examining intra-individual longitudinal stability of cerebrospinal fluid (CSF) biomarkers in Alzheimer's disease, where lumbar puncture is arguably far more invasive than needle arthrocentesis [380, 381, 382, 383]. Consequently, CSF biomarker discovery and therefore validation in Alzheimer's disease is at a more advanced stage than in OA.

One of the major limitations of this study was the small sample size, for reasons already stated, and therefore the lack of statistical power. However, attempts were made to

test the external validity of study patients. It was reassuring that the small study group was similar to a randomly selected group of patients undergoing arthroplasty for end-stage primary knee OA with respect to demographics, radiographic severity, clinical severity (at both time-points) and interval between pre-operative assessment and surgery. The use of the OKS to confirm constancy of clinical severity between time points is biased towards equivalence since the questionnaire asks for responses "during the last 4 weeks" and the median interval was 16.5 days. The VAS_{pain} was, however, is a measure of severity on the day of sampling. The baseline variability of the measured SF cytokines in normal controls over similar short time periods was not tested due to the difficulty in obtaining normal SF. The lack of normal control SF means one cannot completely validate the authenticity of SF concentration measurements from the OA samples. However, it is possible to extrapolate the internal validity of the same SF assays measured in end-stage knee OA samples demonstrated in Chapter 5. The study assessed SF marker temporal variability in end-stage knee OA, where one may expect stability since potential for further deterioration is less than earlier stages. Nonetheless, end-stage knee OA is used to define terminal joint failure in many biomarker studies yet the stability of SF biology has not previously been reported. Finally, this study evaluated the short-term stability of 22 cytokines, chemokines and growth factors that have potential relevance to OA pathophysiology i.e. TGF-β in osteophytosis, but they are by no means exhaustive. The results should not be extrapolated to all prospective SF markers.

The example scenario given in the Introduction of this chapter is clearly unrealistic since it is highly unlikely that a single measurement of any biomarker will determine clinical decisions. A more credible situation would use changes and trends of a combination of biomarkers (biochemical and radiological), and correlation with clinical findings. However, assessment of both short- and long-term stability should be part of the advanced stages of "fit-for-purpose" biomarker validation and should be evaluated at different stages of disease and, if possible, for normal individuals. Any variability, whether systematic or not, should be interpreted in terms of biological and clinical relevance, thus contributing to the concept of minimally important differences for the biomarker in question.

6.6 Conclusions

Null Hypothesis: There is low variability and good absolute agreement between SF analyses made over short periods of time in patients with end-stage knee OA. - ***Accepted***

The study suggests that there is overall stability in the SF biological profile over short intervals during which there has been no change in clinical and radiographic severity. There was some evidence of analyte-specific intra-individual variability, but no directional change between time-points. However, one must be cautious in generalising the results to all possible SF markers. Furthermore, this study was almost certainly underpowered due to difficulties in patient recruitment. Ideally, potential biomarkers highlighted by exploratory studies should be assessed for short-term sampling stability as part of advanced validation. Any variability should be used to define minimally important biological and clinical changes.

This study was presented as a poster at the OARSI World Congress 2013, Philadelphia

Jayadev C, Rout R, Snelling S, Hulley P, Price A. Short-Term Sampling Stability of Synovial Fluid Measurements in Knee Osteoarthritis. Osteoarthr Cartil. 2013 Apr;21:S78–9.

Part II

Wide-spectrum Synovial Fluid Analysis

Chapter 7

Part II Study Design and Methods

7.1 Introduction

Part I of this thesis (Chapters 3 – 6) developed methods and addressed practical implications related to synovial fluid (SF) collection and multiplex analysis. Building on these findings, the focus of the ensuing two chapters is to explore the SF biology of end-stage knee osteoarthritis (OA) with a view to biomarker discovery. This body of work aims to answer two primary questions:

1. Are there differences in the SF biology of structurally progressive phenotypes of knee OA? (Chapter 8)
2. Is there a SF "biological fingerprint" for end-stage knee OA and can it function as a biomarker? (Chapter 9)

This chapter details the study design and methodology used to answer the stated research questions in Part II of the thesis.

7.2 Patient Cohorts

A detailed description of patient groups, including exclusion criteria and ethical considerations are provided in Chapter 2. A summary of patient cohorts involved in the proceeding studies are given in Table 7.1.

Cohort	Description
AMOA	Patients with isolated anteromedial knee OA undergoing either primary unicondylar or total knee arthroplasty
AMOAp	Patients with anteromedial knee OA and patello-femoral osteoarthritis undergoing either primary unicondylar or total knee arthroplasty
LCOA	Patients with lateral compartment knee OA undergoing either primary unicondylar or total knee arthroplasty
TCOA	Patients with tri-compartmental knee OA undergoing primary total knee arthroplasty
Injury	Patients undergoing surgery for cruciate or meniscal injuries without evidence of OA changes (confirmed by radiographs, MRI and at surgery)
Inflammatory	Patients with rheumatoid or psoriatic arthritis of the knee, with or without disease-modifying treatments

Table 7.1 – Description of patient cohorts recruited for studies in Part II of the thesis. The end-stage knee OA (AMOA, AMOAp, LCOA and TCOA) and Injury samples were prospectively collected. Inflammatory samples were retrospectively obtained as a gift and used as "positive controls". Note that collection of Control samples was unsuccessful.

7.3 Patient Assessments

Clinical and structural (radiological and/or intra-operative) assessments were made for patients in the end-stage knee OA and injury cohorts. No assessments were available for patients with retrospectively obtained inflammatory SF samples.

Recruited patients' age, gender and BMI were recorded. Participants were asked to complete the Oxford Knee Score (OKS) questionnaire at pre-operative assessment, or on the morning of surgery if not previously obtained. Patients were assessed with standardised weight-bearing anterior-posterior and lateral views, and a patella skyline view of the index knee conducted at pre-operative assessment clinic. Patients awaiting arthroscopic surgery, also underwent MRI scanning of the index knee within 3 months of surgery. Radiographs were evaluated using the Kellgren and Lawrence (KL) and OARSI Atlas grading systems and MRI scans were assessed qualitatively for the presence of OA features. A detailed description of clinical and radiological assessments is provided in Chapter 2.

7.4 Synovial Fluid Preparation

A detailed description of SF collection and storage is provided in Chapter 2.

Prior to analysis, SF aliquots were thawed at room temperature and clarified at 10000g for 10 minutes. The supernatant was then treated with 2mg/ml bovine testicular hyaluronidase (type I-S, 618.4 U/mg, Sigma). This entailed 1:1 volume mixture of SF with 4mg/ml hyaluronidase, vortexing for 5 seconds and incubation at RT for 1hr on a shaker. Samples were centrifuged at 1000g for 5 minutes and the supernatant used for the assay. The end result was 2-fold sample dilution with 2mg/ml (≈1200U/ml) hyaluronidase. Since SF samples were aliquoted into small volumes at the time of collection, there were no freeze-thaw cycles prior to analysis (at Oxford, see Appendix B.2 on page 284).

7.5 Biological Panel

SF samples were analysed for 35 markers as part of a wide-spectrum panel [Table 7.2].

Pro-inflammatory cytokines	IL-1β, TNF-α, IL-6, IL-8, IL-2, IL-12, IL-15, GM-CSF
Regulatory cytokines	IL-1Ra, IL-4, IL-10, IL-2R
Chemokines	RANTES, MIP-1α, MIP-1β, MCP-1, IP-10, Eotaxin, MIG
Growth Factors	TGF-β1, TGF-β2, TGF-β3, BMP-2, BMP-7
Matrix Enzymes	MMP-1, MMP-3, MMP-9, MMP-13, TIMP-1, ADAMTS-4
Cartilage Turnover	ARGS neo-epitope, COMP, PIIANP
Other (bone)	LIGHT, DcR3

Table 7.2 – Wide-spectrum analysis of synovial fluid: a biological panel of 35 markers

Candidate biomarkers were chosen from soluble mediators and markers previously reported in the literature to be related to OA pathology or pathophysiology. This biological panel is not intended to be definitive nor exhaustive.

7.6 Immunoassays

Based on the findings in Chapters 3 and 4, the Meso Scale Discovery (MSD) platform was favoured over the Luminex platform and magnetic-bead Luminex assays were used in preference in polystyrene-bead assays when possible. However, the choice of immunoassays was largely based on commercial availability, multiplex combinations and ease of developing custom assays if required.

The SF samples in this study were analysed for the 35 markers in the biological panel by 11 different multiplex or singleplex assay kits [Appendix B, Table B.2]. The same platform, assays kit, reagents, lot numbers and protocols were used for each marker throughout the study to analyse all samples. All commercially sourced immunoassays were conducted according to the manufacturer's protocol. Prototype and custom MSD assays were performed using optimised in-house protocols following MSD guidelines and MSD recommended reagents.

On all platforms, calibrators and blanks were measured in duplicate. SF aliquots were run in duplicate for all assays except polystyrene-bead Luminex assays, where they were run in triplicate.

7.7 Censored Data

Concentrations above ULOQ constitute right-censored data, and those below LOQ (or LOD) constitute left-censored data. The impact of data censoring rises with the proportion of censored data. Omission of censored samples could result in considerable data loss and bias in favour of uncensored concentrations. Since different samples may be censored (left or right) for different analytes, data omission may prevent any meaningful correlation or regression analysis. Left censoring is usually more common, and therefore more problematic, than right censoring in biological data since samples are more likely to be below LOQ or LOD than above ULOQ. Analysis of (left) censored data is a large topic and beyond the scope of this thesis [384]. There are several methods to deal with left-censored data including substitutions, log-probit regression, maximum likelihood estimation (MLE) and non-parametric methods

i.e. "reverse" Kaplan-Meier analysis [384, 385, 386, 387]. Most biomarker studies use simple substitutions such as the limit, $^{limit}/_{\sqrt{2}}$, $^{limit}/_{2}$ or even 0. These are easy to perform, but usually introduce significant bias [388]. The other methods above i.e. MLE, which is considered the "gold standard", require familiarity with advanced statistical processes.

Beta-Substitution for Left Censored Data

This study used the β-substitution method for left-censored data with single or multiple censor points (i.e. LOQ and LOD) described by Ganser & Hewett [389]. The distribution of uncensored data is used to calculate a β factor, which is then used to adjust the limit before substitution. The procedure involves separate substitutions to provide summary statistics for the naive data and natural-log transformed data. A sample size of n≥5 is required and data censoring of up to 50% can be handled with performance comparable to the gold standard MLE.

For each analyte, the proportion of samples below LOQ and below LOD in each cohort was recorded. Beta-substitution was performed only if fewer than 50% of samples in the cohort were below LOQ. For each analyte with left-censored data, the β-substitution procedure was conducted separately for patient cohorts, which were treated as distinct data arrays. Samples below LOQ (and above LOD) were substituted with β.LOQ and samples below LOD were substituted with β.LOD. Calculations were made in Microsoft Excel using a template with formulae provided in the original article. If more than 50% of samples in the cohort were below LOQ, then the marker was excluded from further analysis as a continuous variable. However, these markers were retained for qualitative categorical analysis.

Ideally, samples falling above ULOQ for any particular marker should be reanalysed after dilution to bring the concentration within the dynamic range of the assay. However, this was not always possible due to the small sample volumes available and finite funding available for expensive assays. For each marker, the proportion of samples above ULOQ in each cohort was noted. In the unlikely event that there were samples above the dynamic range, concentrations were substituted by the ULOQ for the marker

in question if less than 25% of samples were above ULOQ. If more than 25% of samples in the cohort were >ULOQ, then the marker was excluded from further analysis as a continuous variable.

7.8 Statistical Methods

Markers within the biological panel were classified according to proportion of samples above LOQ to determined subsequent analysis.

- **Quantitative**: Analytes with at least 50% of sample measurements above LOQ in each cohort qualified for quantitative analysis as a continuous variable, following β-substitution of censored data if required.
- **Categorical**: Analytes with fewer than 50% of sample measurements above LOQ in any cohort were excluded from further quantitative concentration analysis. Instead measurements were dichotomised to above/below LOD and above/below LOQ for categorical analysis.

7.8.1 Classical Statistics

The Shapiro-Wilks test was used to assess continuous data for a normal distribution and appropriate transformations were subsequently applied if required. The natural log (to the base e) was used for logarithmic transformations. Specific parametric and non-parametric methods for between-group comparisons were used as appropriate and are stated in the methods sections of the proceeding chapters. In all cases $p<0.05$ was considered significant. Unless otherwise stated, the family-wise error (FWE) from multiple comparisons was corrected using Holm's step-down procedure (Holm-Bonferroni method) [264]. An adjusted p-value (q-value) <0.05 was considered significant.

Statistical analyses and graphing were performed using STATA version 12.1 (StataCorp, College Station, Texas, USA) and PRISM version 5.0 (GraphPad, San Diego, California, USA), respectively.

7.8.2 Multivariate Data Analysis

Classical statistical approaches may not be appropriate for combination biomarker discovery from wide-spectrum sample analysis. Multivariate data analysis (MVDA) extracts information from data with multiple variables by using all the variables simultaneously. MVDA methods are ideally suited to high-dimensional biological data with many more variables than observations. They are robust to collinearity, noise in both X and Y variables and missing data.

Two MVDA approaches were applied to the wide-spectrum SF analysis: principal components analysis (PCA) and partial least squares discriminant analysis (PLS-DA). Both are projection methods that extract information from noise using the correlation structure of the data. They allow powerful graphical data visualisation and simplification using scores that summarise the observations and loadings that summarise the variables. A complete description of these and other MVDA approaches is beyond the scope of this thesis. Therefore readers are directed to relevant introductory texts [390, 391]. In particular, Wold et al. provide an excellent introduction to partial least square (PLS) regression [392].

Data Preparation

SF analyte concentrations were first natural logarithm transformed to minimise data skew. The log transformed variables were then each *auto-scaled* to allow all markers irrespective of range to have equal weight in the analysis [393]. This involved (mean) centering and then scaling to unit variance by dividing by the standard deviation.

Analytes that failed quality control were retained as "categorical markers" and assigned as dummy variables coded 0 for <LOQ and 1 for >LOQ.

Principal Components Analysis (PCA)

PCA was used for initial exploration of SF markers. In the case of categorical analytes, PCA was based on the tetrachoric correlation matrix for binary variables. Principal

components with an eigenvalue below 1 were omitted and varimax orthogonal factor rotation with Kaiser normalisation was applied. Observation score plots were inspected for class separation. Therefore, PCA was conducted as an unsupervised classification method to assess for patient clustering based on SF markers in the biological panel. It was not used as a dimension reduction technique or for determining SF analyte (variable) clustering.

Partial Least Squares Discriminant Analysis (PLS-DA)

If PCA suggested class separation, PLS-DA was conducted to distinguish between patient cohorts on the basis of SF measurements. It was therefore used as a supervised classification method. PLS-DA is an extension of PLS regression when the dependent Y-variable is categorial.

PLS-DA was conducted separately for quantitative and categorical SF analytes and was implemented with the NIPALS algorithm. Predictive models were produced with the study cohort as the categorical dependent Y-variable and the SF analytes as the explanatory X-variables. The R^2 value was used to estimate goodness of model fit i.e. how well the model fits the data. With numerous and correlated X-variables there is a risk for "over-fitting", i.e. getting a well fitting model with little or no predictive power. Therefore a process of internal *cross validation* was conducted to generate a Q^2 value as an estimate of the model's predictive quality i.e. how well the model predicts new data. A total of 7 rounds of cross validation were conducted and a $Q^2 > 0.5$ was considered acceptable. Low R^2 and/or Q^2values indicate that the relationship between X and Y is poor or there is significant noise in the data.

The number of latent projections (components) used in the model was determined by the compromise between optimum R^2 and Q^2 values i.e. the model was stopped at maximum cumulative Q^2 value. Observation score plots were produced to visually assess cohort class separation. Variable loading weights bi-plots were produced to display the relationship between analytes and cohorts. Analytes (X-variables) in the vicinity of a dummy cohort (Y-variable) have the greatest discriminating power.

	Predicted A	**Predicted B**	**Predicted C**	*Accuracy*
Actual A	**TP_A**	$FN_A = FP_B$	$FN_A = FP_C$	$\frac{TP_A}{Total\ Actual\ A}$
Actual B	$FN_B = FP_A$	**TP_B**	$FN_B = FP_C$	$\frac{TP_B}{Total\ Actual\ B}$
Actual C	$FN_C = FP_A$	$FN_C = FP_B$	**TP_C**	$\frac{TP_C}{Total\ Actual\ C}$
Reliability	$\frac{TP_A}{Total\ Predicted\ A}$	$\frac{TP_B}{Total\ Predicted\ B}$	$\frac{TP_C}{Total\ Predicted\ C}$	

Table 7.3 – Confusion matrix for a 3-group predictive model. TP = True Positive, FN = False Negative and FP = False Positive. Please note that false negatives for a particular group represent false positives for another e.g. false negatives for group A (FN_A) also represent false positive for group B (FP_B) and C (FP_C).

Model performance: A confusion matrix was used to assess PLS-DA models by comparing actual cohort to predicted cohort [Table 7.3] . Model accuracy and reliability was calculated for each study cohort. *Accuracy* is the number of patients predicted correctly as a percentage of the total number of patients actually in the cohort. This is equivalent to sensitivity. *Reliability* (or precision) is the number of patients predicted correctly as a percentage of the total number of patients predicted to be in the cohort. This is equivalent to positive predictive value. The average (mean) accuracy and reliability was calculated for each model. The *overall accuracy* of the model was given by the total number of patients correctly classified (i.e. true positives) as a percentage of the total number of patients. *Specificity* is the percentage of patients correctly classified as not belonging to a particular cohort. It was calculated for each cohort by dichotomising the confusion matrix for that cohort e.g. to “group A” and “non-group A”.

SF analyte importance: The variable influence on projection (VIP) parameter (with jack-knifed 95% confidence intervals) is a measure of how much each X-variable (SF analyte) contributes to the overall PLS-DA model. This includes both its importance to class separation (Y-variable) and its importance to modelling the latent structure of X-variables i.e. components. Analytes with a VIP >0.8 were considered important for the overall model; VIP between 0.8 and 0.5 considered potentially important, and VIP <0.5 considered unimportant.

The importance of a given X-variable for Y is proportional to its distance from the origin in the loading space (loading weight bi-plot). These lengths correspond to the PLS regression coefficients, which were therefore used to determine how strongly an analyte is associated with a cohort. The coefficient is significant if its (jack-knifed) 95% confidence interval does not include zero. The "biological fingerprint" of each cohort was defined by its combination of analytes with significant PLS regression coefficients. Markers associated with a cohort (i.e. significant positive coefficient) were termed "positive elements" or "ridges" of the fingerprint; those opposing a cohort (i.e. significant negative coefficient) were termed "negative elements" or "troughs".

Model validation: The data used to generate the PLS-DA models are known as the "training set". Identical wide-spectrum SF analysis of new patients was used as "test set" data to validate the PLS-DA models. The models are blinded to the cohort membership of these new patients and assessed for their ability to correctly classify them. The predictive performance is assessed as described above.

Data processing, PCA and PLS-DA was implemented in SIMCA-P ver. 13.0.2 (Umetrics, Sweden). Coefficient heat maps were created using MeV Ver. 4.8.1 (TG4 Software). Multivariate statistical approaches and results were validated by a statistician (Dr. Gary Collins, Centre for Statistics in Medicine, Oxford) prior to writing the thesis.

Chapter 8

Synovial Fluid Biology of End-stage Knee OA Structural Phenotypes

8.1 Introduction

Osteoarthritis (OA) of the knee is a heterogeneous syndrome that can affect one or more joint compartments. The medial and lateral tibio-femoral and patello-femoral compartments can be involved in isolation or in combination. The medial compartment is most frequently affected with anteromedial OA (AMOA) being the most common phenotype of knee OA [394, 395, 120, 358, 121, 350, 396]. AMOA is a distinct clinical and pathological entity with a highly consistent topographical pattern of tibio-femoral disease [250, 397, 398, 399]. The lateral compartment of the knee has preserved articular cartilage with limited osteophytosis or synovitis. Patello-femoral OA (PFOA) usually occurs in conjunction with tibio-femoral disease, and AMOA plus PFOA represents a significant proportion of bi-compartmental knee OA [120, 358, 121]. Tri-compartmental OA (TCOA) represents a global field change with the disease process affecting all parts of the knee joint.

End-stage knee OA can be defined as clinically and radiologically (structurally) advanced disease, refractory to non-operative management, suitable for knee replacement surgery. In end-stage OA, burden of disease biomarkers should reflect the extent of clinical and structural (radiographic) severity associated with biological joint failure [200]. Consequently, structural phenotypes in end-stage knee OA provide a framework

to discover burden of disease markers since the "volumetric" extent of OA within the knee increases with the number of affected joint compartments.

No previous studies have used well-phenotyped structural patterns of knee OA for synovial fluid (SF) biomarker discovery. Furthermore, there is limited information on the variability or "noise" of SF biomarkers in knee OA. There is evidence from our group that the consistent spatial pattern of disease in AMOA is reflected at a biological level in terms histology [400], chondrocyte gene expression [401], reactive oxygen species and apoptosis [402], and matrix composition [403] .

The highly conserved patho-anatomy and cartilage biology of AMOA may be associated with a less variable SF biology than TCOA, which in contrast is more inconsistent in its distribution and degree of cartilage degeneration.

This chapter explores potential difference in SF cytokines, chemokines, growth factors, matrix enzymes and markers of cartilage turnover in structural phenotypes of end-stage knee OA. The aim of this study was to compare the SF biology, in terms of measured concentrations and variability, in structural phenotypes of end-stage knee OA with cumulative compartment involvement.

8.2 Null Hypotheses

Patients with isolated AMOA, AMOA plus PFOA and TCOA demonstrate equivalent SF biology in terms of:

1. Concentrations of a wide-spectrum of SF markers
2. Variability of measurements for a wide-spectrum of SF markers

Cohort	Description	N
AMOA	Patients with isolated anteromedial knee OA undergoing either primary unicondylar or total knee arthroplasty	15
AMOAp	Patients with anteromedial knee OA and patello-femoral OA undergoing either primary unicondylar or total knee arthroplasty	15
TCOA	Patients with tri-compartmental knee OA undergoing primary total knee arthroplasty	30
Injury	Patients undergoing arthroscopic surgery for cruciate or meniscal injuries without evidence of OA changes	20
Inflammatory	Patients with rheumatoid or psoriatic arthritis of the knee, with or without disease-modifying treatments	18

Table 8.1 – Description of study cohorts and number of synovial fluid samples. The end-stage knee OA (AMOA, AMOAp, LCOA and TCOA) and Injury samples were prospectively collected. Inflammatory samples were retrospectively obtained as a gift and used as "positive controls". Note that collection of Control samples was unsuccessful.

8.3 Methodology

8.3.1 Patient Cohorts

Synovial fluid was collected and analysed from patients with isolated anteromedial osteoarthritis (AMOA), AMOA plus patello-femoral OA (AMOAp), tri-compartmental OA (TCOA), non-osteoarthritic knee injury (Injury) and inflammatory knee arthritis (Inflammatory) [Table 8.1].

For the purpose of testing the hypotheses of this study, SF measurements of knee injury and inflammatory patients were used only to validate SF analysis of the end-stage knee OA cohorts. Any potential differences between the SF biology of knee injury, inflammatory arthritis and end-stage knee OA are explored in Chapter 9 and were therefore not discussed in this chapter.

Larger SF effusion volumes, indicative of greater synovitis, were observed at arthrotomy in patients with TCOA than AMOA or AMOAp. More widespread synovitis was observed intra-operatively during total knee arthroplasty for TCOA than for AMOA. Neither effusion nor synovitis was formally measured.

8.3.2 Clinical and Radiographic Assessments

Patient age was compared by one-way ANOVA with post-hoc Bonferroni multiple comparisons. Between-group comparisons of the Oxford Knee Score (OKS), OKS-Pain and OKS-Function indices were done using the Kruskal-Wallis test with post-hoc Dunn's multiple comparisons. Compartment-specific Kellgren & Lawrence grades (KLG), total OARSI scores, OARSI joint space narrowing (JSN) subtotals and OARSI osteophyte (OP) subtotals are all ordinal in nature. Consequently, between-group comparisons of these radiographic endpoints used the Kruskal-Wallis test with post-hoc Dunn's multiple comparisons. In all cases, $p<0.05$ was considered significant.

8.3.3 Biological Assessment

SF samples were analysed for the biological panel of 35 markers using a combination of Luminex, MSD and ELISA [Table 7.2 on page 185]. Differences between structural phenotypes were first assessed by classical between-group comparisons and then multivariate data analysis.

At least 50% of samples in each cohort were required to be above the limit of quantification (LOQ) for analyte measurements to qualify for quantitative analysis. The β substitution method, described in Chapter 7, was used to handle any (left) censored data and concentrations were treated as a continuous variable. When fewer than 50% of samples in any cohort were above LOQ, the analyte was excluded from further quantitative concentration analysis and was treated as a categorical marker i.e. quantifiable or not.

Heat Maps

Simultaneous graphical representation of analyte concentrations, grouped by cohort, was achieved using heat maps. The natural logarithm of analyte concentrations were standardised using the robust centre scaling method ($x-median/MAD$). Data processing and heat maps production were conducted in Array Studio ver. 5 (Omicsoft Corp., Cary, NC, USA).

Between-group Comparisons

Quantitative analytes: Due to the presence of skewed data, often refractory to logarithmic transformations, the Kruskal-Wallis test with post-hoc Dunn's multiple comparisons was used for comparison of analyte concentrations between end-stage knee OA phenotypes.

Categorical analytes: The frequencies of samples above LOQ (quantifiable) and above LOD (detectable) in each cohort were compared using the chi-squared (χ^2) test (for homogeneity). Due to small sample sizes and the possibility of frequencies less than 5 in the contingency table, a Monte Carlo resampling method (with 5000 simulations) was used to estimate the χ^2 statistic and p-value. If the null hypothesis was rejected, post-hoc pairwise comparisons of proportions between cohorts was conducted using the Marascuilo procedure [404]. The Monte Carlo χ^2 analysis and Marascuilo procedure was implemented in XLSTAT ver. 2013.1.01 (Addinsoft, New York, NY, USA)

Comparison of Cohort Variability

Due to relatively small sample sizes and non-normal data distributions, the variability of clinical, radiographic and biological endpoints for the three end-stage knee OA cohorts was represented by the median absolute deviation from the median (MAD) and robust coefficient of variation (rCV). The rCV was given by MAD/median and expressed as a percentage. This is analogous to standard deviation/mean used to traditionally calculate coefficient of variation. For each endpoint, the variability of the structural phenotypes was compared using the Browne-Forsythe modification of Levene's test (using the median). Post-hoc pairwise comparisons were conducted if the null hypothesis was rejected.

Multivariate Analysis

Concentration measurements for quantitative SF analytes were natural logarithm transformed, (mean) centred and scaled to unit variance. Categorical marker measurements were dichotomised to 0 for <LOQ and 1 for >LOQ.

Exploratory principal components analysis (PCA) was conducted to assess for evidence of class separation on observation score plots. If so, this was followed by supervised classification by partial least squares discriminant analysis (PLS-DA) with end-stage knee OA structural phenotype as the categorical dependent Y-variable and SF analytes as the explanatory X-variables. PCA and PLS-DA were conducted separately for quantitative and categorical analytes. A detailed description of these approaches is provided in Chapter 7 [Section 7.8.2].

8.4 Results

A summary of the clinical and radiographic characteristics of the study cohorts is presented in Table 8.2. Patients in the end-stage knee OA cohorts were well matched for age (~70 years) and BMI (~30 kg/m^2). There was a fairly even split of men and women for patients with TCOA, whereas men predominated for AMOA, with or without PFOA.

The median OKS, OKS-Pain and OKS-Function indices were not significantly different between the end-stage knee OA phenotypes. There was, however, a trend for better scores in patients with AMOAp than TCOA, which had in turn higher OKS and OKS-Pain than those with isolated AMOA.

Compartment specific radiographic KL grading followed the pattern of joint involvement for each cohort. The medial compartment was the most severely affected in all three end-stage knee OA cohorts. Lateral compartment disease was significantly more pronounced in TCOA and not significantly different between AMOA subgroups. The severity of PF joint involvement was equivalent for TCOA and AMOAp, which were both significantly greater than isolated AMOA. The overall severity given by the total

KLG was greatest for patients with TCOA than AMOA with or without PFOA. Although the total KLG was higher in patients with AMOAp than isolated AMOA, this did not reach statistical significance.

The total OARSI score was significantly greater for TCOA than AMOA with or without PFOA, between which there were no significant difference. There was no difference in medial joint space narrowing (JSN) between end-stage knee OA cohorts. However, lateral JSN and therefore overall JSN was significantly greater for TCOA than AMOA subgroups, between which there was no difference. Medial, lateral and overall osteophytosis were significantly more severe for TCOA than AMOA with or without PFOA. Patients with AMOA, irrespective of PFOA, had equivalent medial osteophytosis. Although patients with AMOAp had higher lateral and thus total osteophyte scores than those with isolated AMOA, but this did not reach significance.

8.4.1 Quality Control of SF Analysis

Twenty analytes qualified for quantitative comparisons between cohorts: 5 pro-inflammatory cytokines (TNF-α, IL-6, IL-8, IL-12 and IL-15); 3 chemokines (MCP-1, IP-10 and Eotaxin), 3 isoforms of TGF-β, 5 matrix enzymes (MMP-1, MMP-3, MMP-9, TIMP-1 and ADAMTS-4); 3 markers of cartilage metabolism (374-ARGS neo-epitope, COMP and PIIANP); and DcR3. The frequency of β-substitution was not significantly different between end-stage OA cohort for any of these markers. The average intra-assay precision for each assay was acceptable. All further quantitative analysis of SF concentrations refers to the above markers.

Two of the remaining markers, IL-1β and MMP-13, were quantifiable in over 50% of samples in only the AMOAp cohort. The remaining 13 analytes (IL-2, GM-CSF, IL-1RA, IL-4, IL-10, IL-2R, RANTES, MIP-1α, MIP-1β, MIG, BMP-2, BMP-7 and LIGHT) were not sufficiently quantifiable any end-stage OA cohort. These 15 analytes underwent categorical analysis only.

	AMOA	AMOAp	TCOA	
No. Patients	15	15	30	
Gender	12M : 3F	12M : 3F	14M : 16F	
Age (years)	68.7 (63.2 – 74.2)	70.2 (64.9 – 75.5)	70.6 (67.5 – 73.6)	
BMI (kg/m^2)	28.0 (26.3 – 29.7)	29.3 (27.2 – 31.4)	30.3 (28.1 – 32.58)	
OKS	15 (11 – 27)	25 (15 – 29)	19 (15 – 22.25)	
OKS-Pain	21 (14 – 46)	46 (25 – 54)	36 (24.0 – 43.0)	
OKS-Function	50 (35 – 70)	60 (35 – 70)	45.5 (35 – 55)	
KLG Medial	3 (3 – 3)	3 (3 – 4)	3 (3 – 4)	
KLG Lateral	0 (0 – 1)	1 (0 – 1)	2 (2 – 3)	*
KLG PF	1 (1 – 2)	2 (1 – 2)	2 (2 – 3)	*
KLG Total	4 (3 – 6)	6 (5 – 7)	7 (6.25 – 9.75)	*
OARSI Total	6 (5 – 12)	8 (6 – 11)	14 (11 – 18.25)	*
OARSI JSN	3 (2 – 3)	3 (2 – 3)	4 (3 – 4.25)	*
JSN Medial	3 (2 – 3)	2 (2 – 3)	3 (2 – 3)	
JSN Lateral	0 (0 – 0)	0 (0 – 0)	1 (1 – 2)	*
OARSI OP	3 (2 – 7)	5 (2 – 7)	8.5 (7 – 10)	*
OP Medial	3 (1.5 – 4)	3 (1.5 – 4)	5 (5 – 5.75)	*
OP Lateral	1 (0.5 – 3)	2 (0 – 2.5)	4 (2 – 5)	*

Table 8.2 – Summary of clinical and radiographic assessments for patients with anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp), tri-compartmental OA (TCOA). Radiographic data shown are: Kellgren & Lawrence grades (KLG) for each joint compartment and total KLG score; total OARSI score; OARSI medial and lateral joint space narrowing (JSN) and combined JSN score; OARSI medial and lateral osteophyte (OP) subtotal and combined OP score. The OARSI score does not include assessment of the patello-femoral (PF) compartment. Means (and 95%CI) are given for Age. Medians (and IQR) are given for Oxford Knee Score (OKS); OKS Pain and Function indices; and all radiographic parameters. One-way ANOVA with post-hoc Bonferroni multiple comparisons, or the Kruskall-Wallis test with post-hoc Dunn multiple comparisons were used as appropriate. An asterisk (*) denotes statistical significance ($p<0.05$).

8.4.2 Internal Validity of SF Measurements

SF concentrations for all patients, grouped according to cohort, are graphically represented as a heat map in Figure 8.1. Median analyte concentrations for each cohort are illustrated as a heat map in Figure 8.2. All measurements except ADAMTS-4, 374-ARGS neo-epitope, COMP and PIIANP appear greatest in inflammatory SF samples. COMP and PIIANP levels appear lowest in this cohort. The knee injury cohort appears to have the lowest concentrations for nearly half the analytes (TNF-α, IL-6, IL-8, IL-12, MCP-1, Eotaxin, MMP-1, MMP-3, TIMP-1 and ADAMTS-4) and the highest for 374-ARGS. There was a significant difference across the study cohorts for all quantifiable SF analytes except Eotaxin [Figure 8.2].

For analytes not qualifying for quantitative analysis, the proportions of SF samples above LOQ and LOD, grouped according to cohort, are graphically represented as a heat map in Figure 8.3. There was a significant between-cohort difference in the proportion of quantifiable samples for all analytes except IL-2, IL-4 and BMP-7 [Figure 8.3A]. The proportion of detectable samples was significantly different across cohorts for all markers except the BMPs [Figure 8.3B]. In general, the frequency of quantification and detection was greatest in inflammatory samples and least in injury samples.

8.4.3 Between-group Comparisons of SF Biology

Inflammatory cytokines and DcR3 [Figure 8.4]

The 3 structural phenotypes of end-stage knee OA demonstrated no significant differences in SF concentrations of TNF-α, IL-6, IL-8, IL-12, IL-15 and DcR3. There were a few notable observations and trends that did not reach statistical significance. IL-6 for AMOAp (~120pg/ml) was more than triple that for AMOA and more than double that for TCOA. The concentration of IL-8 in isolated AMOA (~24pg/ml) was approximately half that of AMOAp and TCOA. The median DcR3 concentration increased from AMOA to AMOAp to TCOA, by ~15% at each step.

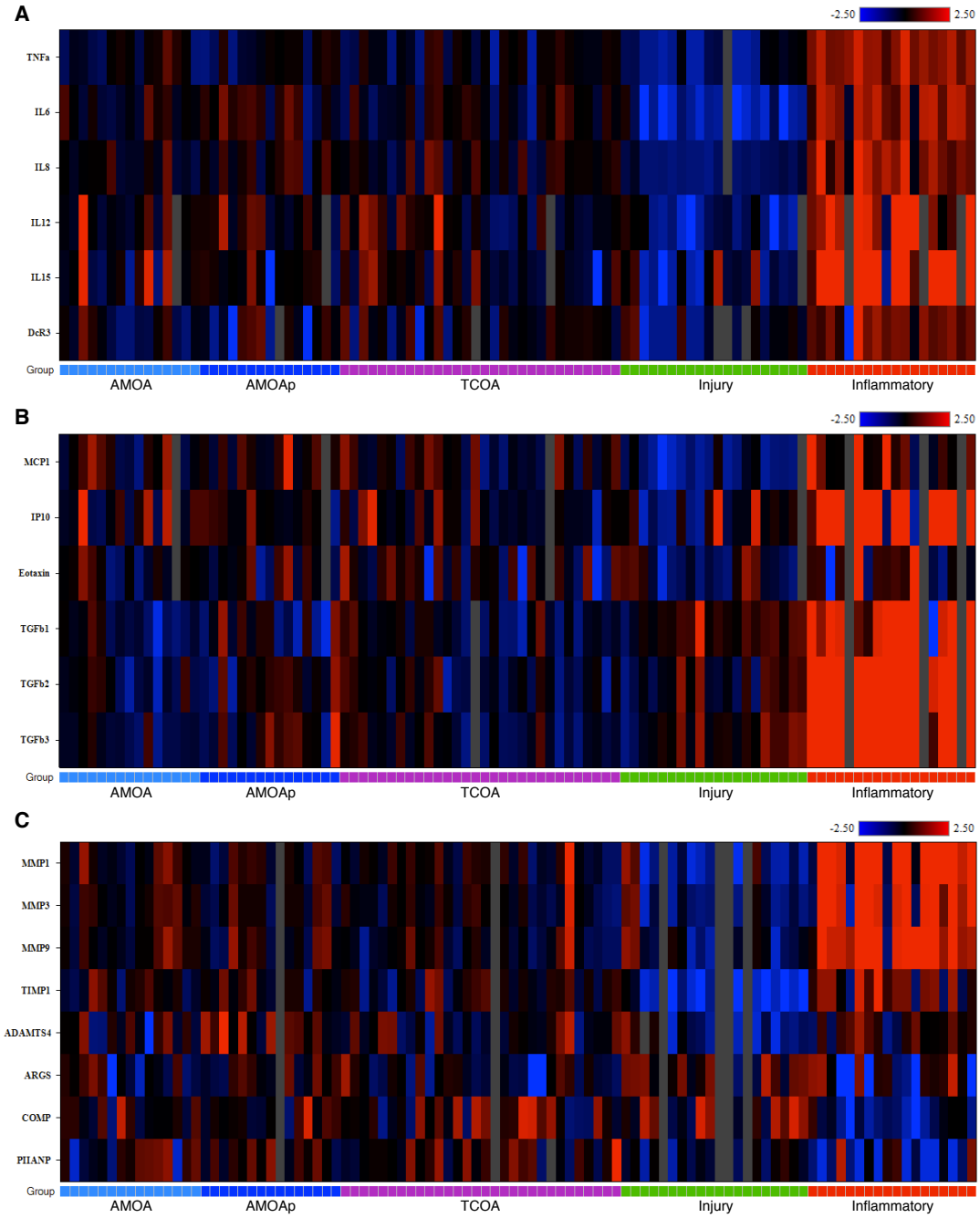


Figure 8.1 – Heat map of relative SF analyte concentrations for individual patients. [A] Pro-inflammatory cytokines and DcR3; **[B]** chemokines and TGF-β isoforms; and **[C]** matrix enzymes and markers of cartilage metabolism. Patients are grouped according to cohort: anteromedial OA (AMOA, light blue), AMOA with patello-femoral OA (AMOAp, dark blue), tri-compartmental OA (TCOA, purple), knee injury (green) and inflammatory arthropathy (red). Analyte concentrations were natural logarithm-transformed then standardised (robust centre scaling). The heat map was generated using Array Studio (ver. 5). Grey boxes indicate missing data.

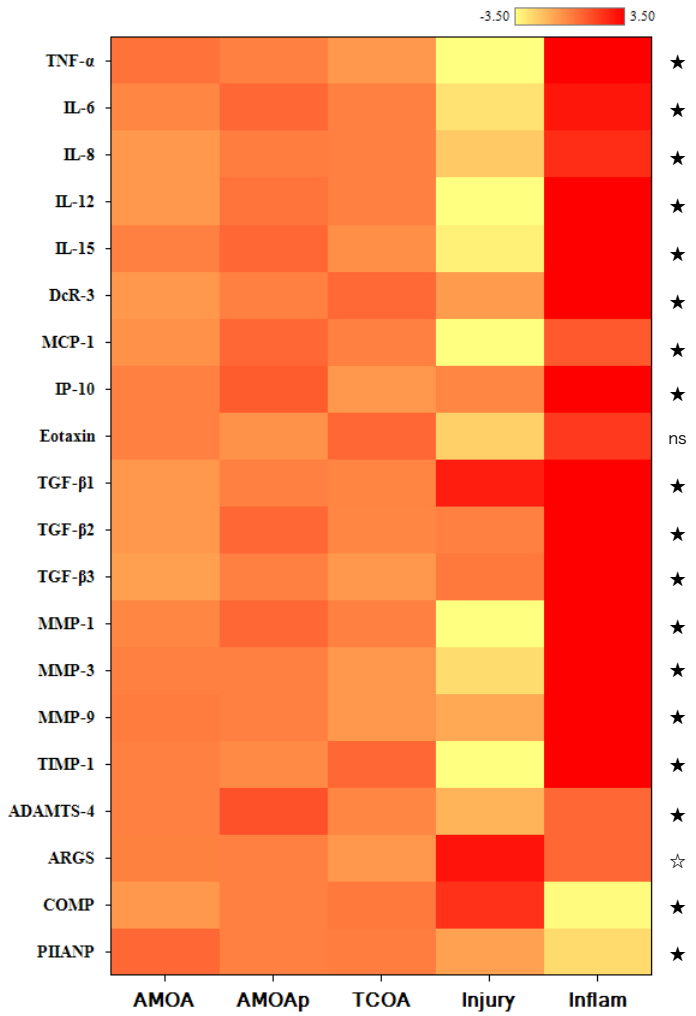


Figure 8.2 – Heat map of standardised median SF analyte concentrations by cohort. Median analyte concentrations for each cohort were natural logarithm transformed and then standardised (robust centre scaling). The heat map was generated using Array Studio (ver. 5). The Kruskal-Wallis test was used to assess differences across cohorts for raw concentrations of each analyte (non-transformed & non-standardised). $P<0.05$ was significant. Stars (★) denote significance ($q<0.05$) after correction for multiple comparisons by Holm's method. Clear stars (☆) denote significance lost after Holm's correction. AMOA – Anteromedial OA; AMOAp – Anteromedial OA with patello-femoral OA; TCOA – tri-compartmental OA; Injury – knee injury; Inflam – inflammatory arthropathy.

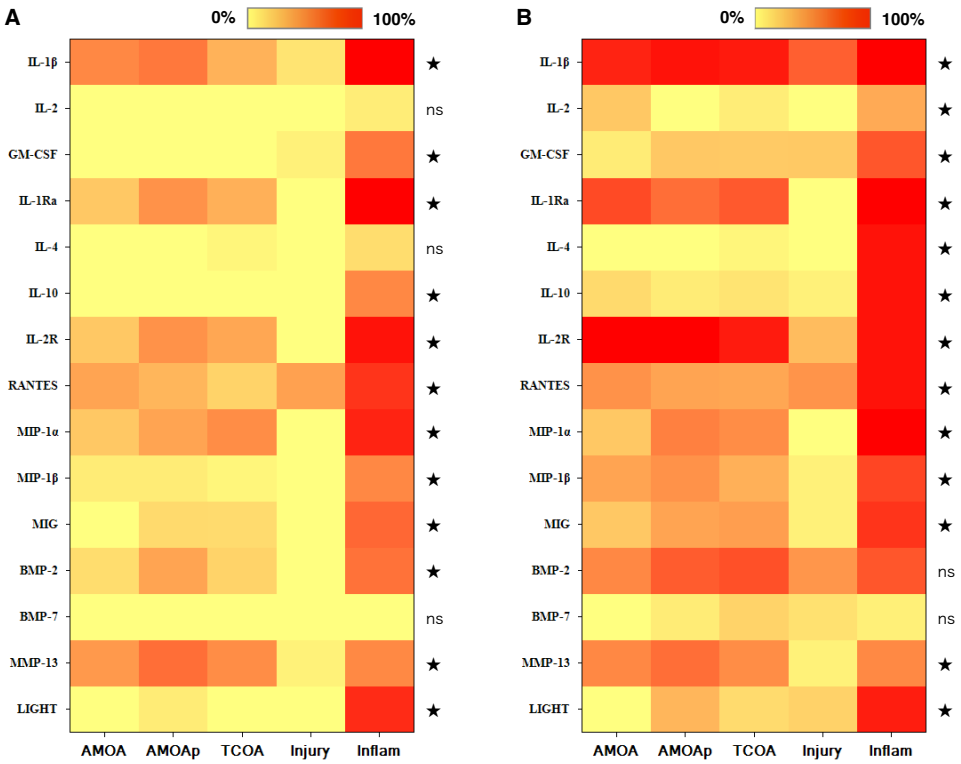


Figure 8.3 – Heat maps for frequency of quantification and detection by cohort for analytes not qualifying for quantitative analysis. [A] Proportion of samples above LOQ. **[B]** Proportion of samples above LOD. Proportions were compared using the χ-squared test (Monte Carlo method) with post-hoc testing by the Marascuilo procedure. $P<0.05$ was significant. Stars (★) denote significance ($q<0.05$) after correction for multiple comparisons by Holm's method. AMOA – Anteromedial OA; AMOAp – Anteromedial OA with patello-femoral OA; TCOA – tri-compartmental OA; Injury – knee injury; Inflam – inflammatory arthropathy.

Chemokines and TGF-β isoforms [Figure 8.5]

There was no significant difference between patients with isolated AMOA, AMOAp and TCOA for SF measurements of MCP-1, IP-10, Eotaxin and the three TGF-β isoforms. Chemokines MCP-1 and IP-10, and TGF-β2 and -3 showed a trend to be highest for AMOAp than isolated AMOA and TCOA.

Matrix enzymes [Figure 8.6]

There was no significant difference in metalloproteinase or aggrecanase concentrations between patients with isolated AMOA, AMOAp and TCOA. However, both MMP-3 and MMP-9 showed a decreasing trend from isolated AMOA to AMOAp to TCOA. The median concentration of MMP-3 was over 370ng/ml (~40%) lower in TCOA than AMOA (with or without PFOA). MMP-9 was about 30% lower in TCOA than AMOA (with or without PFOA). ADAMTS-4 concentration was highest for patients with AMOA plus PFOA; ~17.5μg/ml (40%) higher than isolated AMOA or TCOA. There was no significant difference in TIMP-1 concentration between the structural phenotypes of end-stage knee OA. However, the median concentration was about 10% higher for TCOA than AMOA phenotypes.

Markers of cartilage metabolism [Figure 8.7]

The 3 structural phenotypes of end-stage knee OA demonstrated no significant differences in SF concentrations of 374-ARGS neo-epitope, COMP or PIIANP. However, median concentrations of ARGS neo-epitope showed a trend to be lowest for TCOA. COMP showed an increasing trend from AMOA to AMOAp to TCOA. PIIANP was about 20% greater for isolated AMOA than AMOAp or TCOA.

Categorical Markers

There were no differences in the proportion of samples above LOQ or LOD between the 3 structural phenotypes of end-stage knee OA for the 15 analytes that did not qualify

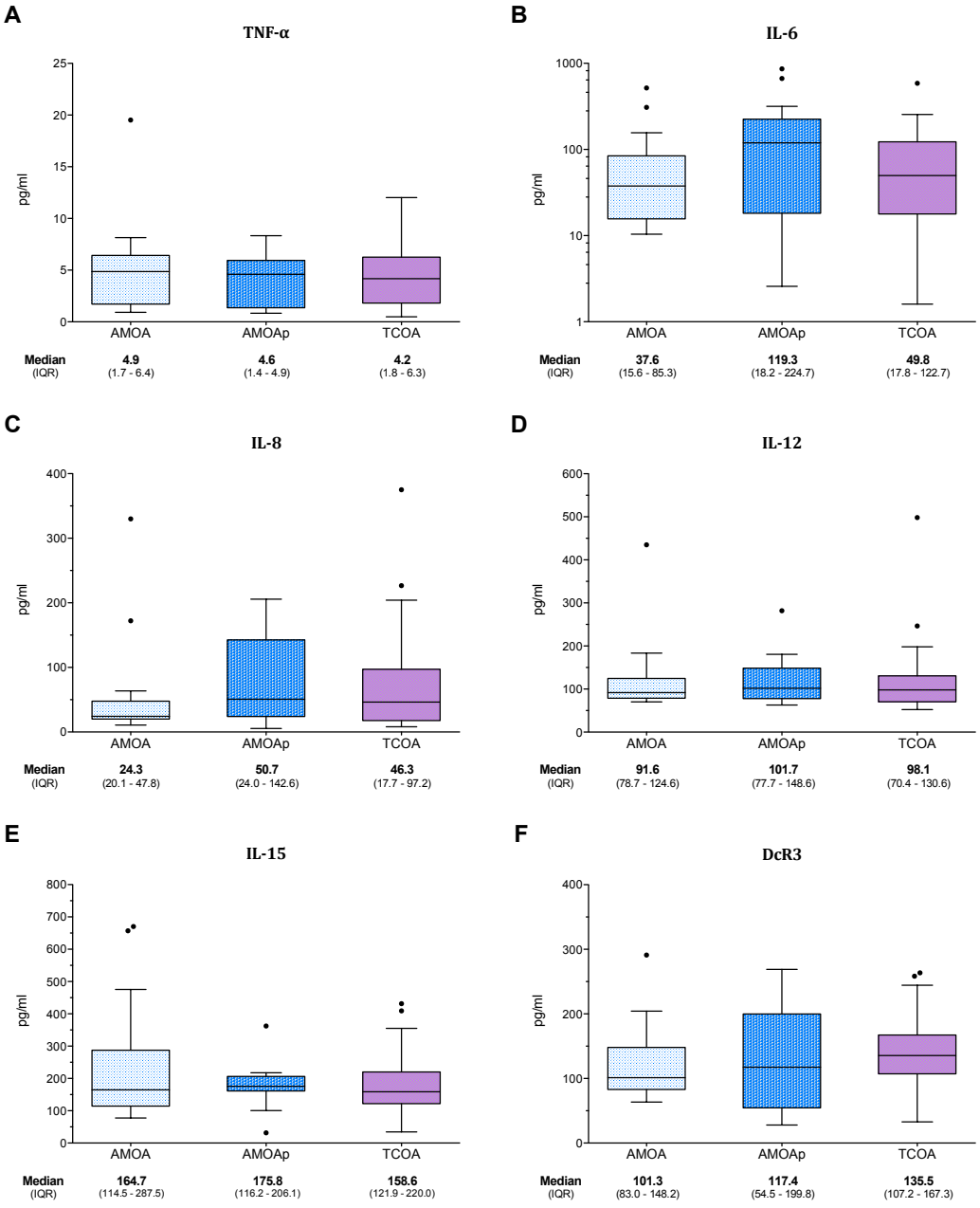


Figure 8.4 – Comparison of pro-inflammatory cytokines and DcR3 between end-stage knee OA phenotypes. [A] TNF-α, **[B]** IL-6, **[C]** IL-8, **[D]** IL-12, **[E]** IL-15 and **[F]** DcR3. Tukey box-and-whiskers are shown for anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp) and tri-compartmental OA (TCOA). Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in pg/ml. Note that IL-6 is given on a logarithmic scale.

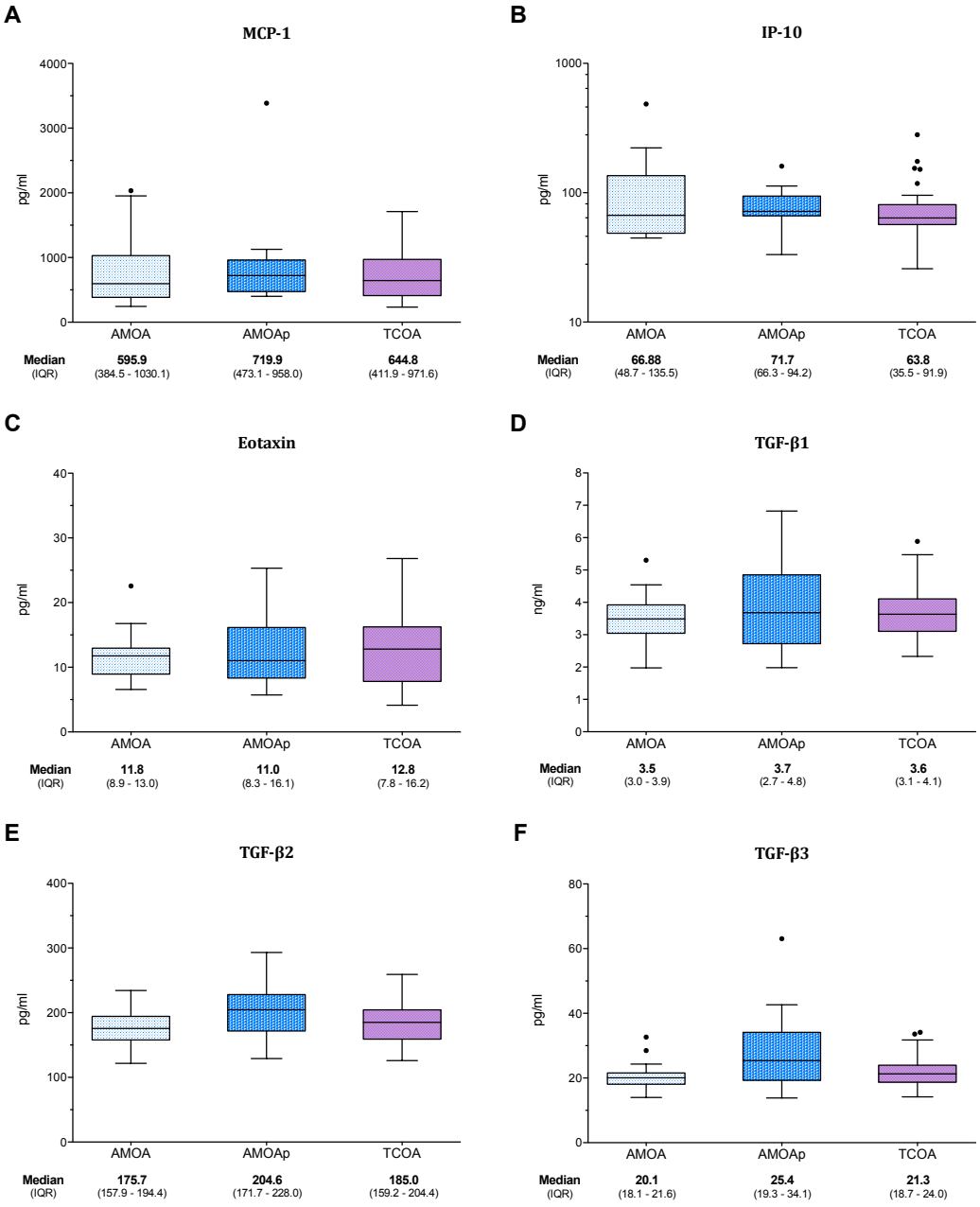


Figure 8.5 – Comparison of chemokines and TGF-β isoforms between end-stage knee OA phenotypes. **[A]** MCP-1, **[B]** IP-10, **[C]** Eotaxin, **[D]** TGF-β1, **[E]** TGF-β2 and **[F]** TGF-β3. Tukey box-and-whiskers are shown for anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp) and tri-compartmental OA (TCOA). Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in pg/ml except TGF-β, which is given in ng/ml. Note that IP-10 is given on a logarithmic scale.

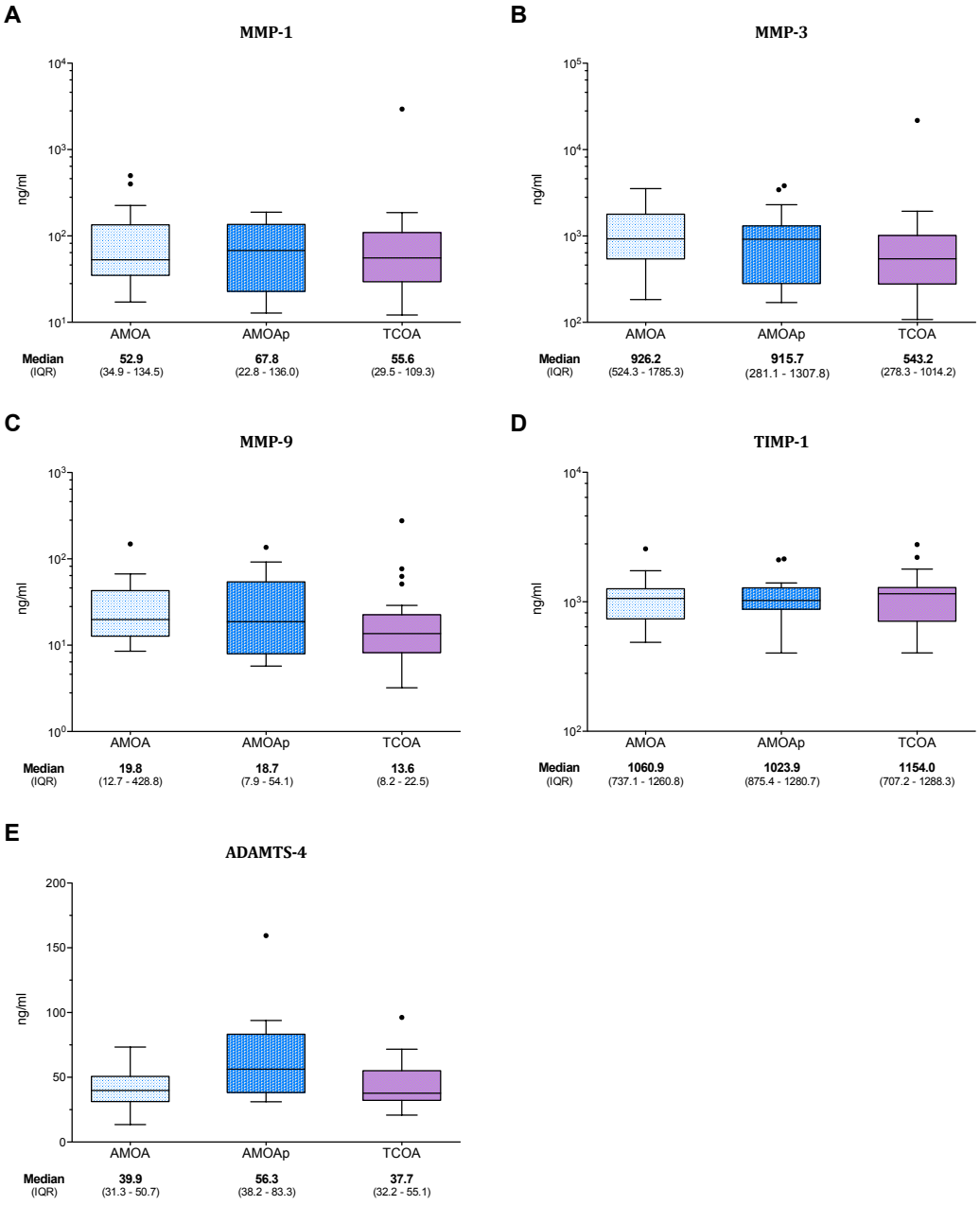


Figure 8.6 – Comparison of matrix enzymes between end-stage knee OA phenotypes. [A] MMP-1, **[B]** MMP-3, **[C]** MMP-9, **[D]** TIMP-1 and **[E]** ADAMTS-4. Tukey box-and-whiskers are shown for anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp) and tri-compartmental OA (TCOA). Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in ng/ml on logarithmic scales except ADAMTS-4, which is shown on a linear scale.

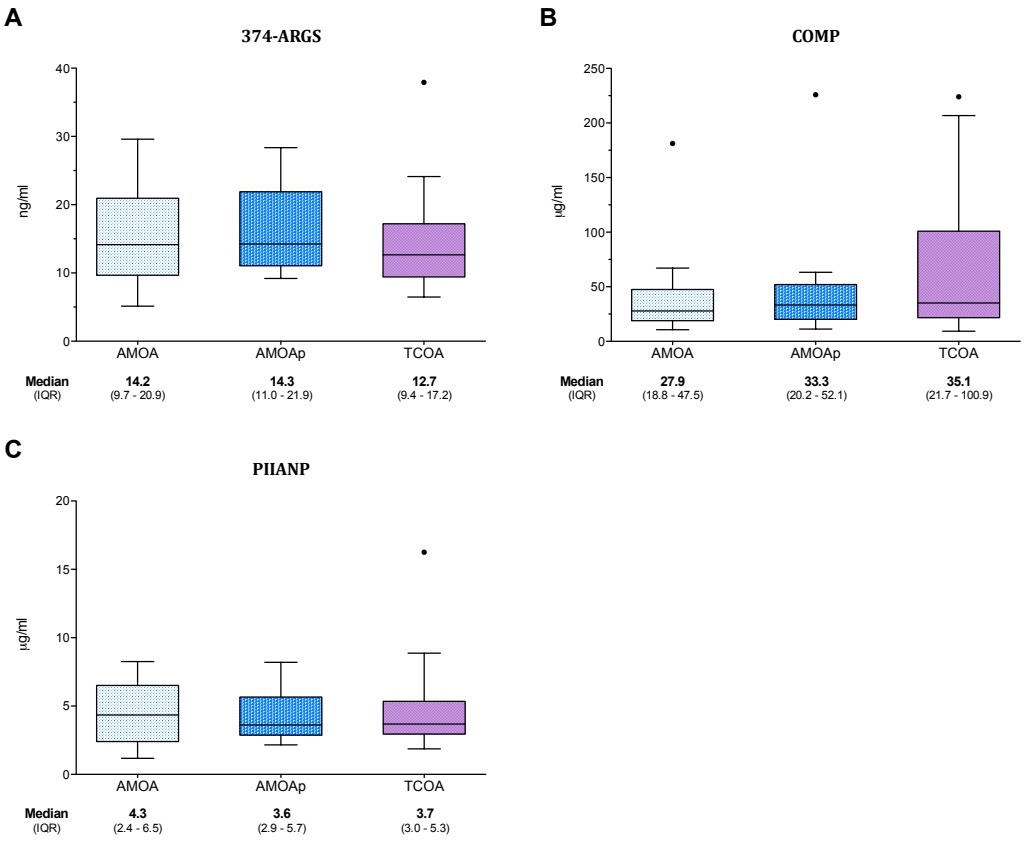


Figure 8.7 – Comparison of markers of cartilage metabolism between cohorts. [A] 374-ARGS aggrecan neo-epitope, **[B]** COMP and **[C]** PIIANP. Tukey box-and-whiskers are shown for anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp) and tri-compartmental OA (TCOA). Measurements outside the Tukey whiskers are shown as black dots. COMP and PIIANP are given in µg/ml and 374-ARGS is given in ng/ml.

	AMOA	AMOAp	TCOA	p-value
IL-1β	46.7%	53.3%	30%	0.318
IL-2	0%	0%	0%	-
GM-CSF	0%	0%	0%	-
IL-1Ra	21.4%	42.9%	31.0%	0.428
IL-4	0%	0%	3.4%	1.000
IL-10	0%	0%	0%	-
IL-2R	21.4%	42.9%	34.5%	0.509
RANTES	35.7%	28.6%	17.2%	0.360
MIP-1α	21.4%	35.7%	44.8%	0.369
MIP-1β	7.1%	7.1%	3.4%	1.000
MIG	0%	14.3%	13.8%	0.512
BMP-2	13.3%	35.7%	17.2%	0.287
BMP-7	0%	0%	0%	-
MMP-13	40%	57.1%	44.8%	0.694
LIGHT	0%	7.1%	0%	0.202

Table 8.3 – Proportion of SF samples above LOQ for analytes not qualified for quantitative concentration analysis. The proportion of samples above LOQ is given as a percentage for each cohort: isolated anteromedial OA (AMOA, n=15), AMOA plus patello-femoral OA (AMOAp, n=15) and tri-compartmental OA (TCOA, n=30). Proportions were compared using the χ-squared test (Monte Carlo method) with post-hoc testing by the Marascuilo procedure. $P<0.05$ was considered statistically significant.

for quantitative concentrations comparisons [Table 8.3 and Table 8.4]. There was, however, a trend for several categorical analytes to be quantifiable and/or detectable in a greater proportion of AMOAp samples than isolated AMOA and occasionally TCOA.

8.4.4 Variability of SF biology

The results of variability analysis between the 3 structural phenotypes of end-stage knee OA are summarised in Table 8.5. Patients with isolated AMOA, AMOAp and TCOA demonstrated no difference in variability for either clinical scores (OKS, OKS-Pain and OKS-Function) or radiographic scores (KLG total, OARSI total, OARSI joint space narrowing and OARSI osteophyte).

The 3 structural phenotypes of end-stage knee OA demonstrated no significant differences in variability for SF measurements of 18 (out of the 20) analytes: 5 pro-inflammatory cytokines (TNF-α, IL-6, IL-8, IL-12 and IL-15); 3 chemokines (MCP-1, IP-10 and Eotaxin); TGF-β2; 5 matrix enzymes (MMP-1, MMP-3, MMP-9, TIMP-1

	AMOA	AMOAp	TCOA	p-value
IL-1β	86.7%	93.3%	90%	1.000
IL-2	21.4%	0%	6.9%	0.052
GM-CSF	7.1%	21.4%	20.7%	0.545
IL-1Ra	71.4%	57.1%	65.5%	0.718
IL-4	0%	0%	3.4%	1.000
IL-10	14.3%	7.1%	10.3%	1.000
IL-2R	100%	100%	89.7%	0.413
RANTES	42.9%	35.7%	34.5%	0.096
MIP-1α	21.4%	50%	44.8%	0.230
MIP-1β	35.7%	42.9%	31.0%	0.757
MIG	21.4%	35.7%	37.9%	0.583
BMP-2	46.7%	64.3%	69%	0.399
BMP-7	0%	7.1%	17.2%	0.136
MMP-13	46.7%	57.1%	44.8%	0.886
LIGHT	0%	28.6%	13.8%	0.084

Table 8.4 – Proportion of SF samples above LOD for analytes not qualified for categorical concentration analysis. The proportion of samples above LOD is given as a percentage for each cohort: isolated anteromedial OA (AMOA, n=15), AMOA plus patello-femoral OA (AMOAp, n=15) and tri-compartmental OA (TCOA, n=30). Proportions were compared using the χ-squared test (Monte Carlo method) with post-hoc testing by the Marascuilo procedure. $P<0.05$ was considered statistically significant.

and ADAMTS-4); 3 markers of cartilage metabolism (374-ARGS neo-epitope, COMP and PIIANP); and DcR3. TGF-β1 ($p=0.023$) and TGF-β3 ($p=0.013$) did not exhibit homogeneity of variance between end-stage knee OA cohorts. Both were more variable in patients with AMOAp than isolated AMOA or TCOA, between which there was no significant difference.

8.4.5 Multivariate Analysis of SF Biology

Multivariate analysis was conducted for 20 quantitative and 8 categorical SF analytes [Table 8.6]. Quantitative markers included 5 inflammatory cytokines (TNF-α, IL-6, IL-8, IL-12 and IL-15); 3 chemokines (MCP-1, IP-10 and Eotaxin), 3 isoforms of TGF-β, 5 matrix enzymes (MMP-1, MMP-3, MMP-9, TIMP-1 and ADAMTS-4); 3 markers of cartilage metabolism (374-ARGS neo-epitope, COMP and PIIANP); and DcR3. Categorical markers included inflammatory cytokine IL-1β; 2 regulatory cytokines (IL-1Ra and IL-2R); 3 chemokines (RANTES, MIP-1α and MIG); growth factor BMP-2 and matrix enzyme MMP-13.

	AMOA		AMOAp		TCOA		
	MAD	rCV	MAD	rCV	MAD	rCV	
OKS	6.00	40%	8.00	32%	4.00	21%	
OKS-Pain	7.14	33%	11.47	25%	7.14	20%	
OKS-Function	15.00	30%	15.00	25%	9.68	21%	
KLG Total	1.00	25%	1.00	17%	1.00	14%	
OARSI Total	1.00	17%	2.00	25%	3.00	21%	
OARSI JSN	1.00	33%	0.00	0%	1.00	25%	
OARSI OP	1.00	33%	2.00	40%	1.50	18%	
TNF-α	2.08	43%	1.96	43%	2.12	51%	
IL-6	23.04	61%	101.16	85%	42.53	85%	
IL-8	10.10	42%	31.66	62%	30.76	66%	
IL-12	14.36	16%	24.92	24%	29.52	30%	
IL-15	53.42	32%	20.98	12%	39.08	25%	
DcR3	37.77	37%	62.95	54%	31.24	23%	
MCP-1	223.06	37%	240.12	33%	236.52	37%	
IP-10	19.56	29%	9.08	13%	11.40	18%	
Eotaxin	1.88	16%	2.76	25%	3.84	30%	
TGF-β1	433.38	12%	954.98	26%	387.32	11%	*
TGF-β2	17.81	10%	29.88	15%	20.39	11%	
TGF-β3	1.68	8%	7.83	31%	2.60	12%	*
MMP-1	24.41	46%	47.19	70%	28.59	51%	
MMP-3	519.86	56%	630.26	69%	304.71	56%	
MMP-9	7.53	38%	11.64	62%	5.81	43%	
TIMP-1	312.44	29%	172.35	17%	309.73	27%	
ADAMTS-4	10.350	26%	18.78	33%	6.57	17%	
ARGS	4.47	32%	4.59	32%	3.01	26%	
COMP	11940.00	43%	13170.00	40%	22600.00	64%	
PIIANP	2014.10	46%	924.75	26%	1126.95	31%	

Table 8.5 – Variability of clinical, radiographic and biological endpoints for structural phenotypes of end-stage knee OA. The median absolute deviation from the median (MAD) and robust coefficient of variation (rCV = MAD/median) is given for each endpoint. Analyte MADs are given in pg/ml except for matrix enzymes (MMP-1, MMP-3, MMP-9, TIMP-1 and ADAMTS-4) and markers of cartilage turnover (374-ARGS neo-epitope, COMP and PIIANP), which are given in ng/ml. The Browne-Forsythe modification of Levene's test (using the median) was used to assess homogeneity of variance between isolated anteromedial OA (AMOA), AMOA plus patello-femoral OA (AMOAp) and tri-compartmental OA (TCOA). An asterisk (*) denotes statistical significance ($p<0.05$).

	Quantitative	Categorical
Inflammatory cytokines	TNF-α, IL-6, IL-8, IL-12, IL-15	IL-1β
Regulatory cytokines		IL-1Ra, IL-2R
Chemokines	MCP-1, IP-10, Eotaxin	RANTES, MIP-1α, MIG
Growth Factors	TGF-β1, TGF-β2, TGF-β3	BMP-2
Matrix Enzymes	MMP-1, MMP-3, MMP-9, ADAMTS-4, TIMP-1	MMP-13
Cartilage Turnover	374-ARGS, COMP, PIIANP	
Others (bone)	DcR3	

Table 8.6 – Quantitative and Categorical markers used for multivariate analysis

IL-2, GM-CSF, IL-10 and BMP-7 were excluded because they were not quantifiable in any end-stage knee OA phenotypes [Table 8.3]. IL-4, MIP-1β and LIGHT were also excluded as they were >LOQ in only one sample in one or more cohort.

Score plots from PCA of quantitative or categorical markers failed to show any differentiation between end-stage knee OA structural phenotypes [Figure 8.8]. Therefore, further assessment by PLS-DA was abandoned.

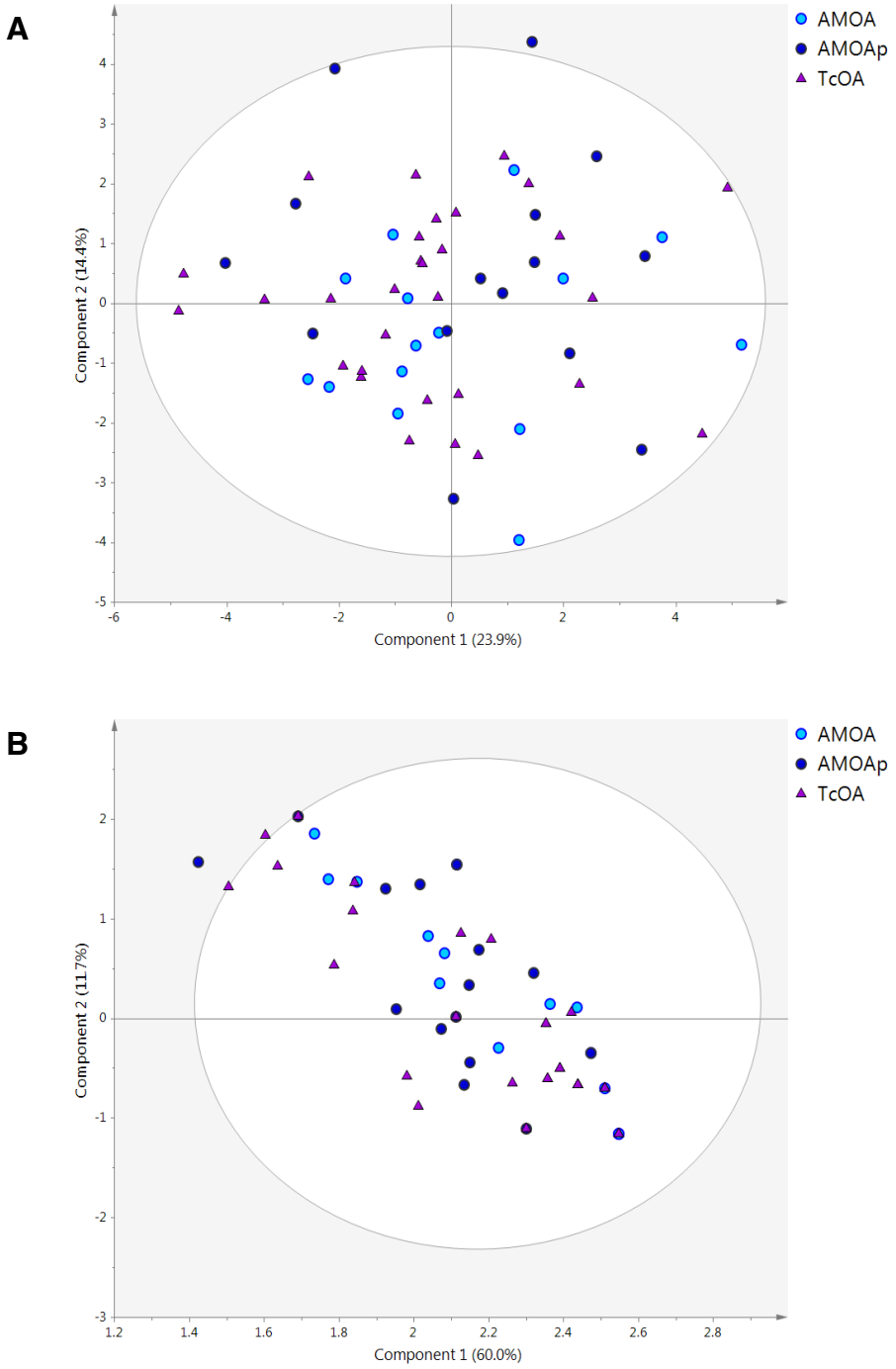


Figure 8.8 – PCA observation score plots for end-stage knee OA structural phenotypes. [A] quantitative and **[B]** categorical SF analytes did not demonstrate any obvious class separation between patients with anteromedial OA (AMOA), AMOA with patello-femoral OA (AMOAp) or tri-compartmental OA (TCOA). There were 6 principal components with eigenvalues >1 for both quantitative and categorical marker analysis. The first 2 components are shown on the score plots .

8.5 Discussion

This study aimed to assess whether the SF biology of end-stage knee OA reflects the increasing volume of pathology from unicompartmental to tricompartmental joint involvement, thereby revealing potential burden of disease biomarkers. In addition, it assessed if the variability in SF biology of anteromedial knee OA mirrored its highly conserved structural and biological pattern of disease. Despite significant differences in radiographic burden of disease, the results demonstrate no differences in either the concentration or variability of a wide range of cytokines, chemokines, growth factors, matrix enzymes and markers of matrix turnover between structural phenotypes of end-stage knee OA.

This study took a novel and alternative approach to biomarker discovery using well-characterised phenotypes of end-stage knee OA classified according to structural patterns of increasing compartmental joint involvement. All patients were well matched for age, BMI and clinical severity. Radiographic assessments, confirmed intra-operatively, verified differences in structural burden of disease with progressive compartment involvement. Although both standard comparison and multivariate approaches yielded no potential biomarker, the results pose a very important and challenging problem. Why does focal OA confined to one joint compartment have an equivalent SF biology to extensive 3-compartment degeneration for a comprehensive range of markers implicated in OA pathophysiology?

The validity of SF immunoassay measurements in the end-stage knee OA phenotypes is not in doubt since significant differences were observed for almost all markers when compared to patients with non-osteoarthritic knee injuries and inflammatory arthritis. These findings will be explored further in Chapter 9. Only the chemokine Eotaxin and growth factor BMP-7 showed poor internal validity. The latter may have been due to inadequate analytical assay sensitivity. There was strict consistency in methodology to facilitate evaluation and comparison of data. For each marker assayed, all patient samples underwent identical pre-analytical and analytical steps as advocated by the findings in Chapters 3 and 4. Furthermore post-analytical data processing and quality control was also standardised. Potential bias could have been introduced by

β-substitution of left censored data. A high proportion of censored data in a group may mask differences since true concentration values could be in a range of values less than the substitution value. However, the proportion of samples requiring β-substitution was not significantly different between end-stage OA cohorts for any analyte. Although the frequency of quantification or detection was equivalent between structural phenotypes for markers failing quantitative analysis, it remains possible that more sensitive assays could reveal phenotypic differences.

SF was sampled from the most severely affected, and largely equivalently affected, medial compartment in all 3 structural phenotypes. However, medial compartment SF sampling is unlikely to be the reason for the lack of phenotypic differences. This was highlighted in Chapter 4 that demonstrated that the anatomical site of sampling does not affect SF analysis even when there are significant compartmental differences in structural burden of disease. It is acknowledged that only 22 of the present 35 markers were assessed in the previous study, and in particular matrix enzymes and markers of cartilage turnover were not evaluated. Nonetheless, the significant multi-collinearity of biological systems makes compartment-specific biology for these additional markers unlikely. Thus far, one can reasonably conclude that the comparable SF analyte concentrations between structurally progressive patterns of end-stage knee OA are real and not an analytical artefact or the result of "physiological compartmentation" of the synovial space.

At advanced stages of disease, it is plausible that macroscopically preserved joint compartments in AMOA suffer pathological changes at a microscopic level. Moen et al. demonstrate histological OA changes (according to the Mankin grading system) in the radiographically normal lateral compartment of patients with end-stage AMOA irrespective of PFOA [405]. Other studies have reported that radiographically and macroscopically normal lateral cartilage in patients with end-stage AMOA is mechanically inferior and has biological features of OA [406, 407]. To the contrary, Brocklehurst et al. demonstrated no biochemical or histological differences between macroscopically preserved femoral condyle cartilage in patients with end-stage knee OA and in age-matched controls [408]. By extrapolation of the former studies, the PF compartment of patients with isolated end-stage AMOA in this study could also exhibit

pre-macroscopic OA changes. One can speculate that the true burden of disease in AMOA (with or without PFOA) is greater than radiographically graded resulting in higher concentrations of SF markers than expected. This mirrors the current paradigm for OA pathogenesis in which the development of joint failure long pre-dates the appearance of macroscopic changes. The evidence available suggests, at most, pre-macroscopic OA in preserved compartments cannot fully explain why SF markers levels in single-, bi- and tri-compartmental OA were equivalent. One can venture two possible explanations that are not mutually exclusive: a "ceiling effect" in which measurable SF biological activity reaches a saturation point at a far earlier stage of disease; and a "burnout phenomenon" in which SF biological activity falls as progressively more articular volume is lost with cumulative compartment involvement. The latter may account for the trend in MMP-3, MMP-9 and ARGS neo-epitope to be lower in TCOA than AMOA. Benito et al. also demonstrated higher mononuclear cell infiltrates and over expression of inflammatory mediators in the synovium of patients with early knee OA compared to end-stage knee OA [81]. Longitudinal natural history studies are required to conclusively test both these concepts.

The biological relevance of concentration measurements must be interpreted with caution. The lack of significant differences in marker concentrations between structural phenotypes does not necessarily imply "equivalent SF biology". The concentration of a molecule is a static measure that is used as a surrogate for the rate of release (or production), but it is equally dependent on the rate of egress (or clearance) from the joint compartment [208]. Differences (or equivalence) in concentrations are often (mis)interpreted in terms of production since it is assumed that clearance is uniform amongst all tested subjects or time-points. This is not the case. For example, SF protein clearance rates in rheumatoid arthritic knees is on average twice that in OA knees, which in turn has accelerated clearance compared to control knees [409, 410]. Clearance of SF molecules occurs at the synovium; either via lymphatic drainage, or direct diffusion into capillaries (for small molecules i.e. cytokines). Even low-grade synovitis significantly increases SF protein clearance [411]. TCOA is associated with widespread synovitis, whereas focal medial peri-meniscal synovitis is observed in 50% of patients with AMOA [83, 412]. In this study, patients with TCOA had larger effusions and more widespread synovial changes indicative of synovitis than patients

with AMOA (with or without PFOA). Unfortunately, neither parameter was formally measured. Greater synovitis and thus higher SF protein clearance rates in TCOA could give concentration measurements equivalent to AMOA even though the marker of interest is being produced at a significantly higher rate in TCOA. Although SF marker concentrations between AMOA and TCOA are the same, the underlying biological activity related to burden of disease may be very different.

There was a trend for several markers to be higher in AMOAp than isolated AMOA or TCOA. This could also be explained in terms of synovitis and SF kinetics. Despite pathology in an additional joint compartment, AMOAp may not have proportionately more synovitis than isolated AMOA. The greater production of markers, reflecting greater disease burden, is not offset by appreciably greater synovial clearance and hence marker concentrations were somewhat higher. As explained above, TCOA has greater synovitis and clearance rates commensurate with further structural involvement that negates higher marker production due to disease burden.

The variability of markers in AMOA did not reflect its highly repeatable topographical pattern of disease and cartilage biology. This may be a statistical artefact of a small (and uneven) samples size susceptible to outliers. Factors that could contribute to increased biological variability in AMOA include activity, analgesia i.e. non-steroidal anti-inflammatory drugs, mechanical alignment, degree of synovitis, and extent of pre-macroscopic disease in preserved cartilage. However, all of these factors apply equally to the other structural phenotypes in the study. It is more likely that (knee) OA remains a highly heterogeneous syndrome despite best efforts of classification through strict inclusion and exclusion criteria, and clinical and structural characterisation.

There were shortcomings to this study and some have already been mentioned above. The number of patients was relatively small and there were unequal group sizes with the two AMOA groups having half the number as the TCOA group. There was a comparatively large number of analytes (v) in relation to the number of patients (n). This has consequences for standard statistical analysis. However, efforts were taken to adjust for multiple comparisons and multivariate approaches such as PCA are robust to unfavourable n/v ratios. Arguably the biggest limitation of this study is that the extent and distribution of synovitis was not evaluated. Phenotyping of end-stage knee OA

patients was based on radiographic scores that are used as a proxy for OA severity, and are weighted for (inferred) cartilage loss, bone remodelling and osteophytosis. There is now overwhelming evidence for the role of synovium in the pathophysiology of OA [reviewed in 75, 77, 39]. Furthermore, as discussed above, synovitis can have a huge impact on SF concentration measurements due its influence on SF protein clearance rates. However, accurate assessment of synovitis would require either intra-operative arthroscopy or pre-operative MRI to allow formal grading, and should be considered for future studies.

8.6 Conclusions

Null Hypotheses: Patients with isolated AMOA, AMOA plus PFOA and TCOA demonstrate equivalent SF biology in terms of:

1. Concentrations of a wide-spectrum of SF markers. ***- Accepted***
2. Variability of measurements for a wide-spectrum of SF markers. ***- Accepted***

A novel approach using cumulatively more severe structural phenotypes of end-stage knee OA failed to reveal burden of disease SF biomarkers. There were no phenotypic differences in the concentration of a wide range of cytokines, chemokines, growth factors, matrix enzymes and markers of matrix turnover. Pre-macroscopic OA changes in preserved cartilage could have increased disease activity within the joint and the SF biology may have plateaued at an earlier stage of progression ("ceiling phenomenon"). End-stage TCOA may also represent a burnt-out joint with reduced biological activity. Alternatively, the findings may be a paradoxical consequence of unaccounted synovitis associated with more extensive structural disease. The principles of SF kinetics and the central role of synovitis in SF kinetics are well established, but underappreciated. They have significant implications for the interpretation of SF concentration measurements and therefore biomarker research. Whatever the true explanation, the findings suggest that patients with end-stage knee OA have equivalent clinical severity and a common SF profile reflecting a terminal state joint failure irrespective of structural involvement.

Chapter 9

Synovial Fluid Fingerprinting of End-stage Knee Osteoarthritis

9.1 Introduction

Although the literature has produced many potential candidate biomarkers in osteoarthritis (OA), there are presently no reliable, quantifiable and easily measured biomarkers that enable diagnosis, prognosis or monitoring of treatment effect at the individual level. OA is a heterogeneous disease of the entire synovial joint with a complex myriad of pathological processes. Therefore, it is not entirely surprising that the pursuit of single biomarkers has thus far proved largely unsatisfactory. A multi-marker approach comprising a profile of several combined biomarkers is likely to be more appropriate [238].

According to the current paradigm, end-stage OA is analogous to end-stage organ failure requiring transplant. End-stage knee OA can be defined as clinically severe and radiologically advanced disease, refractory to non-operative management, suitable for replacement surgery. It also represents the only known true convergence of clinical symptoms and structural pathology in OA. Synovial fluid (SF) bathes the articular space and acts as a medium of communication for the triangular relationship between synovium, cartilage and bone that is central to the pathophysiology of OA [413, 39]. There is a vast body of literature implicating soluble mediators such as cytokines, chemokines, growth factors and matrix enzymes in the loss of homeostasis between

synthetic and degradative processes that ultimately leads to biological and mechanical failure of the joint [Table 1.4 on page 22]. Furthermore, the by-products of cartilage matrix destruction and attempted repair are released into the articular space. The SF composition in end-stage OA should therefore reflect the pathophysiology of this terminal state of "organ" failure. The identification of a signature or "fingerprint" for end-stage knee OA from wide-spectrum SF analysis would provide a portrait of joint death that is easily recognised and, therefore, valuable to monitor disease progression and assess interventions. Consequently, such a SF fingerprint would represent a combination biomarker that qualifies for overlapping BIPEDS classifications. Since it is linked to a definite clinical endpoint i.e. failed non-operative treatment defined clinically and structurally with the need for arthroplasty, the SF fingerprint of end-stage knee OA could also signify a true surrogate endpoint in OA research; the "virtual knee replacement".

Although the interest in systems biology "-omics" approaches to OA biomarker discovery is gathering pace, the technology is expensive and presently not accessible to many laboratories. There have been no previous studies applying systems biology concepts to accessible wide-spectrum immunoassay analysis for OA biomarker discovery. The previous chapter demonstrated that patients with end-stage knee OA have equivalent clinical severity and a common SF profile reflecting a terminal state of joint failure irrespective of structural involvement. This final experimental chapter addresses the ultimate aim of the thesis to discover a SF biological fingerprint for end-stage knee OA from a wide array of pathophysiologically relevant cytokines, chemokines, growth factors, matrix enzymes and markers of cartilage turnover.

9.2 Null Hypothesis

Wide-spectrum SF analysis cannot provide a combination biomarker ("biological fingerprint") for end-stage knee OA.

Cohort	Description	N
esOA	Patients with clinically and radiologically advanced disease, refractory to non-operative management, undergoing primary knee arthroplasty. This included patients with anteromedial OA (with or without PFOA) and tri-compartmental knee OA	60
Injury	Patients undergoing arthroscopic surgery for cruciate or meniscal injuries without evidence of OA changes (confirmed by MRI and intra-operatively)	20
Inflammatory	Patients with rheumatoid or psoriatic arthritis of the knee, with or without disease-modifying treatments	18
Validation	Patients with lateral compartment knee OA undergoing either primary unicondylar or total knee arthroplasty	10

Table 9.1 – Description of study cohorts and number of synovial fluid samples. The end-stage knee OA (esOA), Injury and Validation samples were prospectively collected. Inflammatory samples were retrospectively obtained as a gift. The relatively low patient numbers in the validation cohort reflect the lower prevalence of isolated primary lateral compartment OA. Note that collection of Control samples was unsuccessful.

9.3 Methodology

9.3.1 Patient Cohorts

Clinical, radiographic and biological data from patients previously presented in Chapter 8 were used. Patients with anteromedial OA, with or without patello-femoral OA, and tri-compartmental OA were grouped into a single end-stage knee OA (esOA) cohort. Patients with non-osteoarthritic knee injury (Injury) and inflammatory knee arthritis (Inflammatory) were used as reference groups.

Additional SF samples from patients with lateral compartment OA undergoing arthroplasty were prospectively collected and analysed. These patients provided a validation test group for the multivariate models (see below). The study groups are summarised in Table 9.1.

9.3.2 Clinical and Radiographic Assessments

Patient age was compared by one-way ANOVA with post-hoc Bonferroni multiple comparisons. Between-group comparisons of the Oxford Knee Score (OKS), OKS-Pain and OKS-Function indices were done using the Kruskal-Wallis test with post-hoc Dunn's

multiple comparisons. Radiographic parameters included the overall (worst) Kellgren & Lawrence grade (KLG), total of compartment-specific KLG, total OARSI score, OARSI joint space narrowing (JSN) subscore and OARSI osteophyte (OP) subscore. Between-group comparisons of these radiographic endpoints used the Kruskal-Wallis test with post-hoc Dunn's multiple comparisons. There was no clinical or radiographic data for the inflammatory cohort, as samples were obtained retrospectively from another study. In all cases, $p<0.05$ was considered significant.

9.3.3 Biological Assessment

SF samples were analysed for a wide-spectrum panel of 35 markers using a combination of Luminex, MSD and ELISA [Table 7.2 on page 185]. Differences between structural phenotypes were first assessed by classical between-group comparisons and then multivariate data analysis.

At least 50% of samples in each cohort were required to be above the limit of quantification (LOQ) for analyte measurements to qualify for quantitative analysis. The β-substitution method, described in Chapter 7, was used to handle any (left) censored data and concentrations were treated as a continuous variable. When fewer than 50% of samples in any cohort were above LOQ, the analyte was excluded from further quantitative concentration analysis and was treated as a categorical marker i.e. quantifiable or not.

Between-group Comparisons

Quantitative analytes: Due to the presence of skewed data, often refractory to logarithmic transformations, the Kruskal-Wallis test with post-hoc Dunn's multiple comparisons was used for comparison of analyte concentrations between study cohorts.

Categorical analytes: The frequencies of samples above LOQ (quantifiable) and above LOD (detectable) in each cohort were compared using the chi-squared (χ^2) test (for homogeneity). Due to small sample sizes and the possibility of frequencies less than

5 in the contingency table, a Monte Carlo resampling method (with 5000 simulations) was used to estimate the χ^2 statistic and p-value. If the null hypothesis was rejected, post-hoc pairwise comparisons of proportions between cohorts was conducted using the Marascuilo procedure [404]. The Monte Carlo χ^2 analysis and Marascuilo procedure was implemented in XLSTAT ver. 2013.1.01 (Addinsoft, New York, NY, USA)

Multivariate Analysis

A detailed description of the multivariate approach is provided in Chapter 7 [Section 7.8.2]. Concentration measurements for quantitative SF analytes were natural logarithm transformed, (mean) centred and scaled to unit variance. Categorical marker measurements were dichotomised to 0 for <LOQ and 1 for >LOQ. Exploratory principal components analysis (PCA) was conducted to assess for evidence of class separation on observation score plots. If so, this was followed by supervised classification by partial least squares discriminant analysis (PLS-DA) with study cohort as the categorical dependent Y-variable and SF analytes as the explanatory X-variables. PCA and PLS-DA were conducted separately for quantitative and categorical analytes.

Model performance was assessed by a confusion matrix and ability to diagnose end-stage knee OA was validated using 10 patients with end-stage (lateral compartment) knee OA that were used as a test cohort. The "biological fingerprint" of each cohort was defined by its combination of analytes with significant standardised PLS regression coefficients.

The PLS-DA process was repeated to obtain a streamlined model with the most parsimonious combination of quantitative SF markers for class discrimination. An iterative approach was used to obtain the greatest R^2 and Q^2 values with the least number of quantitative markers that all had a VIP>0.5.

9.4 Results

A summary of the clinical and radiographic characteristics of the study cohorts is presented in Table 9.2. The three groups of patients had significantly different ages.

	esOA	Injury		Inflammatory	
No. Patients	60	20		18	
Gender	38M : 22F	16M : 4F		7M : 11F	
Age	70.0 (67.2 – 72.3)	26.0 (22.9 – 29.1)		48.4 (37.7 – 59.2)	*
OKS	19 (14.25 – 25.5)	40 (35.75 – 42.0)	*		
OKS-Pain	36 (21 – 46)	79 (68.75 – 86.0)	*		
OKS-Function	50 (35 – 64)	90 (85.75 – 94.75)	*		
KLG Worst	3 (3 – 4)	0 (0 - 0)	*		
KLG Total	6 (5 – 8)	0 (0 - 0)	*		
OARSI Total	11 (7 – 14)	0 (0 - 0)	*		
OARSI JSN	3 (2 – 4)	0 (0 - 0)	*		
OARSI OP	7 (3.5 – 9)	0 (0 - 0)	*		

Table 9.2 – Clinical and radiographic assessments for study cohorts. No clinical or radiographic assessments were available for the inflammatory cohort because samples were obtained as a gift. Means (and 95%CI) are given for Age. Medians (and IQR) are given for Oxford Knee Score (OKS); OKS Pain and Function indices; and all radiographic parameters. Radiographic data shown are: Kellgren & Lawrence grades (KLG) in the worst compartment and total KLG score; total OARSI score; OARSI joint space narrowing (JSN) subtotal; and OARSI combined (OP) subtotal. One-way ANOVA with post-hoc Bonferroni multiple comparisons, or the Kruskall-Wallis test with post-hoc Dunn multiple comparisons were used as appropriate. An asterisk (*) denotes statistical significance ($p<0.05$).

Those with knee injuries (~25 years) were the youngest, and the inflammatory cohort (~50 years) were aged approximately midway between the end-stage knee OA and injury patients. Men predominated in the end-stage knee OA and injury cohorts, whereas women predominated in the inflammatory cohort.

The median OKS, OKS-Pain and OKS-Function indices were significantly better (higher) in patients with knee injuries than the end-stage knee OA cohorts. By definition, patients with knee injury have no radiological features of OA so all radiographic scores were significantly higher in the end-stage knee OA cohort.

9.4.1 Quality Control of SF Analysis

Twenty analytes qualified for quantitative comparisons between cohorts: 5 pro-inflammatory cytokines (TNF-α, IL-6, IL-8, IL-12 and IL-15); 3 chemokines (MCP-1, IP-10 and Eotaxin), 3 isoforms of TGF-β, 5 matrix enzymes (MMP-1, MMP-3, MMP-9, TIMP-1 and ADAMTS-

4); 3 markers of cartilage metabolism (374-ARGS neo-epitope, COMP and PIIANP); and DcR3. The average intra-assay precision for each assay was acceptable. All further quantitative analysis of SF concentrations refers to the above markers.

Nine of the remaining markers (IL-1β, GM-CSF, IL-1RA, IL-2R, RANTES, MIP-1α, MIG, BMP-2 and LIGHT) were sufficiently quantifiable in only inflammatory samples. The final 6 analytes (IL-2, IL-4, IL-10, MIP-1β, MMP-13 and BMP-7) were not quantifiable in sufficient numbers of patients in any cohort. These 15 analytes underwent categorical analysis only.

9.4.2 Between-group Comparisons of SF Biology

Inflammatory cytokines and DcR3 [Figure 9.1]

Concentrations of TNF-α, IL-6, IL-8, IL-12, IL-15 and DcR3 were greatest in inflammatory SF than the end-stage knee OA or knee injury samples ($p<0.001$). Patients with knee injuries had a lower SF concentration of TNF-α, IL-6, IL-8 and IL-12 than patients with end-stage knee OA ($p<0.001$). The median SF concentration of IL-15 was significantly lower (~25% less) in knee injury than the end-stage OA cohorts ($p<0.033$), but this lost significance after correction for multiple comparisons. Although DcR3 was ~25% higher in end-stage knee OA than knee injury, this did not reach significance.

Chemokines and TGF-β isoforms [Figure 9.2]

Synovial fluid MCP-1 concentration was least in patients with knee injury ($p<0.001$), and there was no significant difference between the end-stage knee OA and inflammatory cohorts. The IP-10 concentration was greatest in the inflammatory cohort ($p<0.001$), and there was no significant difference between patients with knee injury or end-stage knee OA. Eotaxin levels were not significantly different between cohorts.

SF concentrations of both TGF-β1 and TGF-β3 were lowest in the end-stage knee OA cohort. There was no difference in TGF-β2 level between patients with knee injury and end-stage knee OA. The concentrations of all three TGF-β isoforms were greatest in

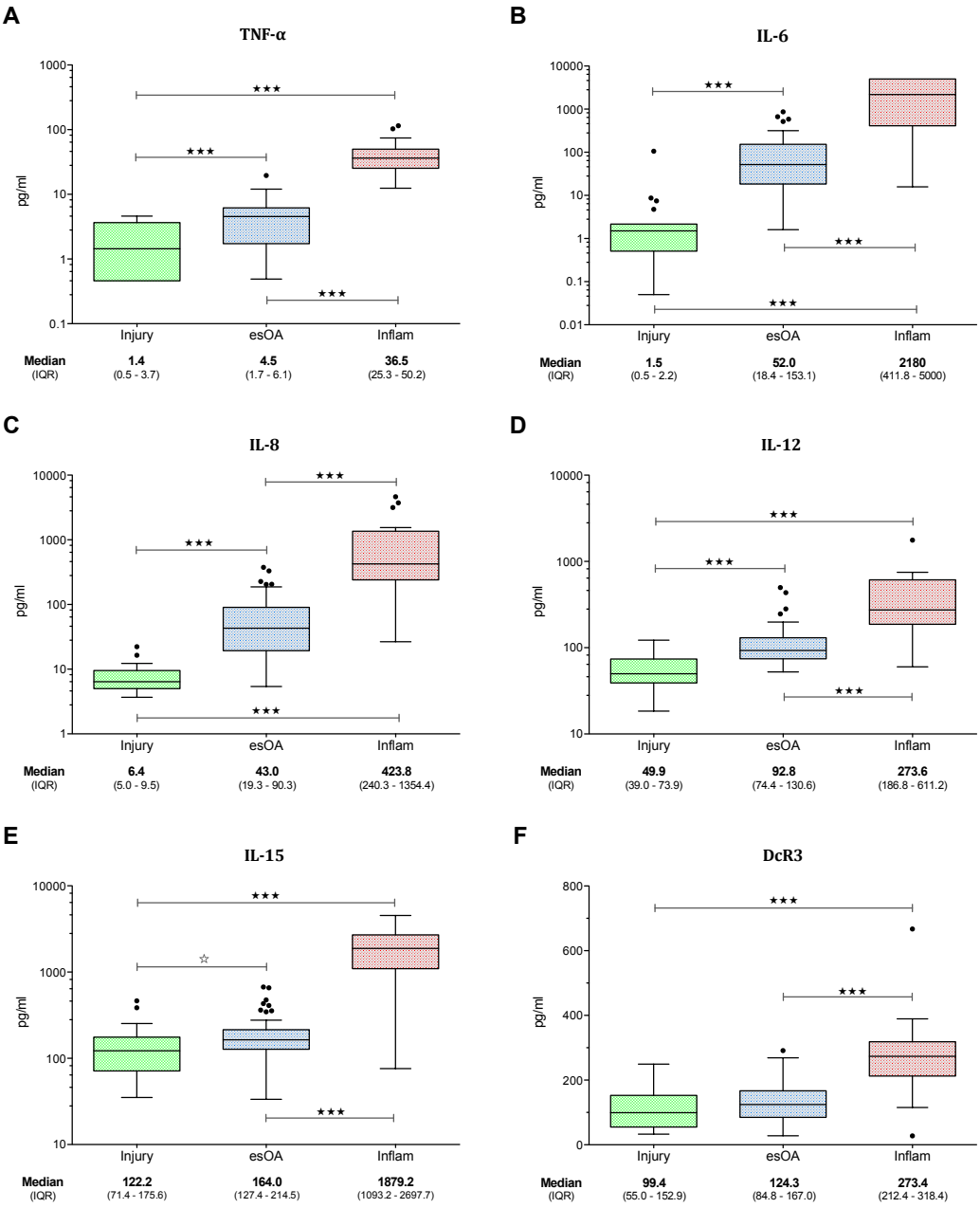


Figure 9.1 – Comparison of pro-inflammatory cytokines and DcR3 between cohorts. [A] TNF-α, **[B]** IL-6, **[C]** IL-8, **[D]** IL-12, **[E]** IL-15 and **[F]** DcR3. Tukey box-and-whiskers are shown for knee injury (Injury), end-stage knee OA (esOA) and inflammatory (Inflam) cohorts. Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in pg/ml on logarithmic scales except DcR3 that is given on a linear scale. Statistical significance is denoted by ★ ($p<0.05$), ★★ ($p<0.01$) and ★★★ ($p<0.001$). Clear symbol (i.e. ☆) denotes significance lost after Bonferroni correction for multiple comparisons.

inflammatory SF samples, but the difference in TGF-β1 between the inflammatory and injury cohorts lost significance after correction of multiple comparisons.

Matrix enzymes [Figure 9.3]

Synovial fluid MMP-1 and MMP-3 were greater in patients with end-stage knee OA than knee injury ($p<0.01$). Although the median SF concentration of MMP-9 was >50% greater in end-stage knee OA than knee injury, this did not reach significance. SF concentrations of MMP-1, MMP-3 and MMP-9 were greatest in the inflammatory cohort ($p<0.01$). The ADAMTS-4 concentration was lowest for patients with knee injury ($p<0.01$), but there was no significant difference between end-stage knee OA and inflammatory cohorts. The concentration of metalloproteinase inhibitor, TIMP-1, was least in patients with knee injury ($p<0.001$). Although TIMP-1 in inflammatory samples was ~800ng/ml (~75%) greater than the end-stage knee OA cohorts ($p<0.05$), this lost significance after correction for multiple comparisons.

Markers of cartilage metabolism [Figure 9.4]

The median SF concentration of 374-ARGS neo-epitope was lowest in end-stage knee OA patients, but was only significant when compared to the knee injury cohort that has nearly double the concentration. Although levels in the knee injury cohort were 50% higher than the inflammatory cohort, this did not reach significance. COMP was higher in knee injury and end-stage knee OA than the inflammatory cohorts ($p<0.001$). Knee injury patients had the highest COMP levels being ~1.5 times that of the end-stage knee OA patients, but this did not reach significance. PIIANP concentrations were highest in the end-stage knee OA cohorts; being about double that of inflammatory patients ($p<0.001$) and ~30% higher than knee injury patients ($p<0.05$). However, the difference between knee injury and end-stage knee OA cohorts was lost after correction for multiple comparisons. SF PIIANP was ~50% higher in knee injury than inflammatory patients, but this did not reach significance.

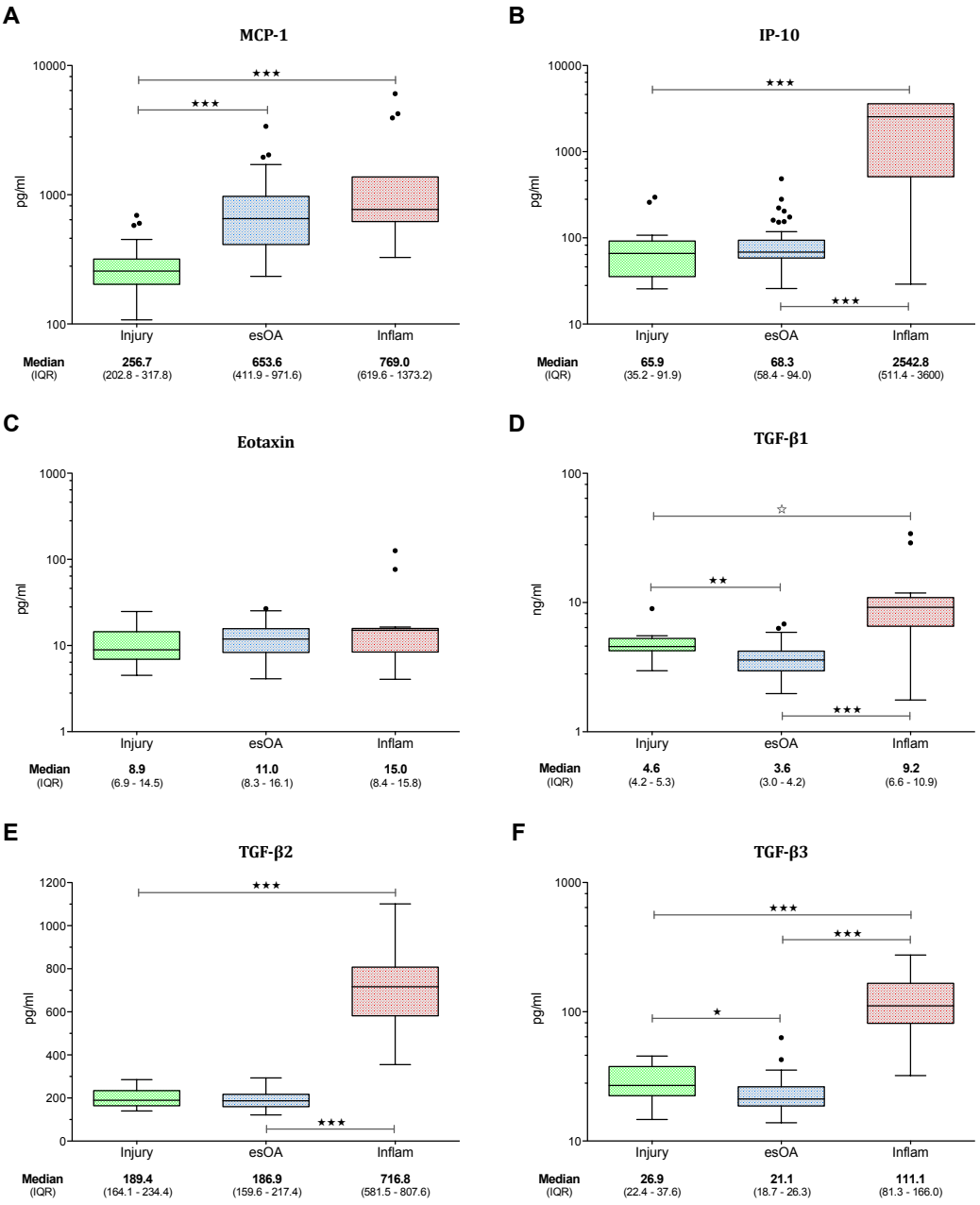


Figure 9.2 – Comparison of chemokines and TGF-β isoforms between cohorts. [A] MCP-1, **[B]** IP-10, **[C]** Eotaxin, **[D]** TGF-β1, **[E]** TGF-β2 and **[F]** TGF-β3. Tukey box-and-whiskers are shown for knee injury (Injury), end-stage knee OA (esOA) and inflammatory (Inflam) cohorts. Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in pg/ml except TGF-β1 that is given in ng/ml. Logarithmic scales are used except for TGF-β2 that is given on a linear scale. Statistical significance is denoted by ★ ($p<0.05$), ★★ ($p<0.01$) and ★★★ ($p<0.001$). Clear symbol (i.e. ☆) denotes significance lost after Bonferroni correction for multiple comparisons.

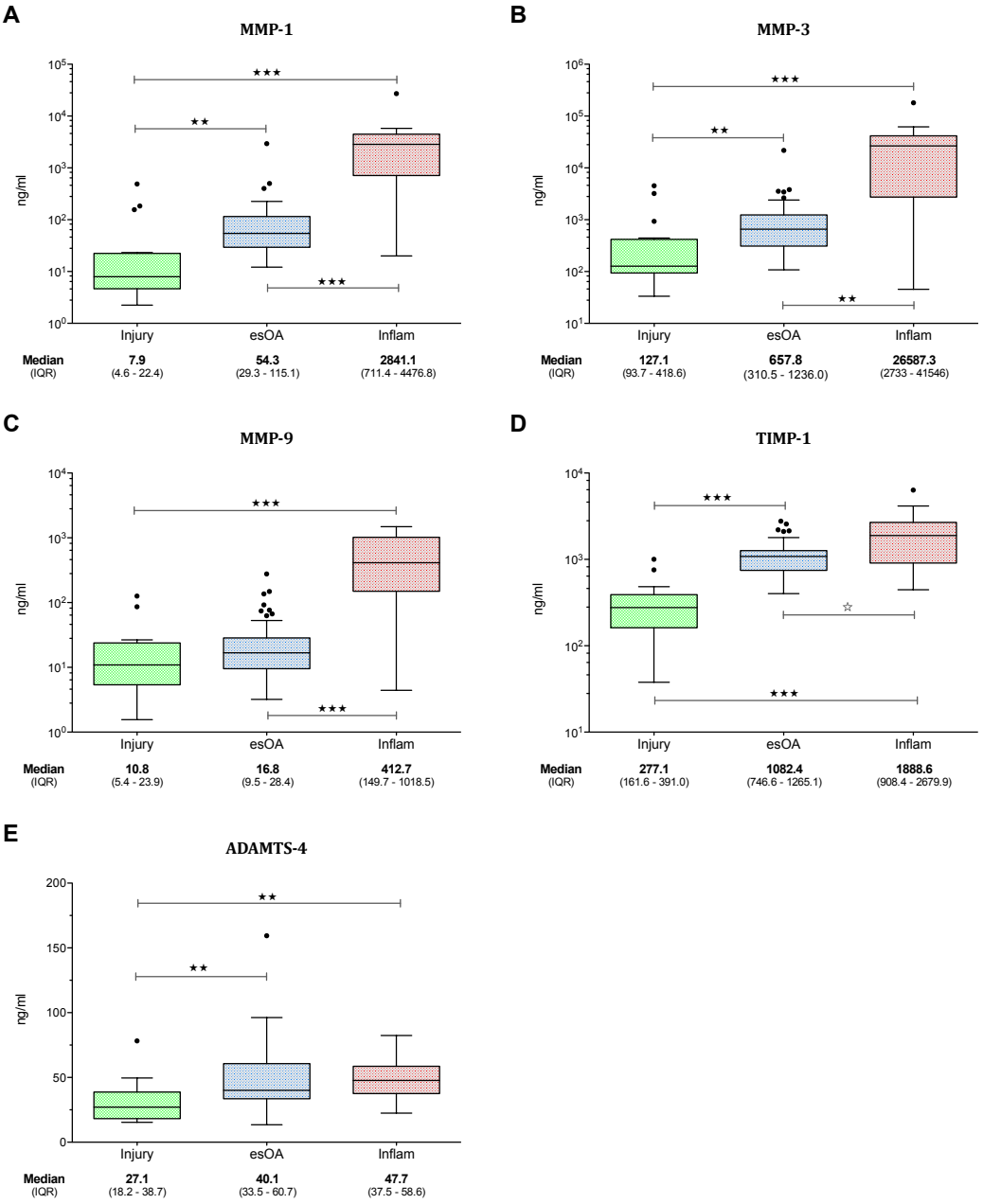


Figure 9.3 – Comparison of matrix enzymes between cohorts. [A] MMP-1, **[B]** MMP-3, **[C]** MMP-9, **[D]** TIMP-1 and **[E]** ADAMTS-4. Tukey box-and-whiskers are shown for knee injury (Injury), end-stage knee OA (esOA) and inflammatory (Inflam) cohorts. Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in ng/ml on logarithmic scales except ADAMTS-4 that is given on a linear scale. Statistical significance is denoted by ★ ($p<0.05$), ★★ ($p<0.01$) and ★★★ ($p<0.001$). Clear symbol (i.e. ☆) denotes significance lost after Bonferroni correction for multiple comparisons.

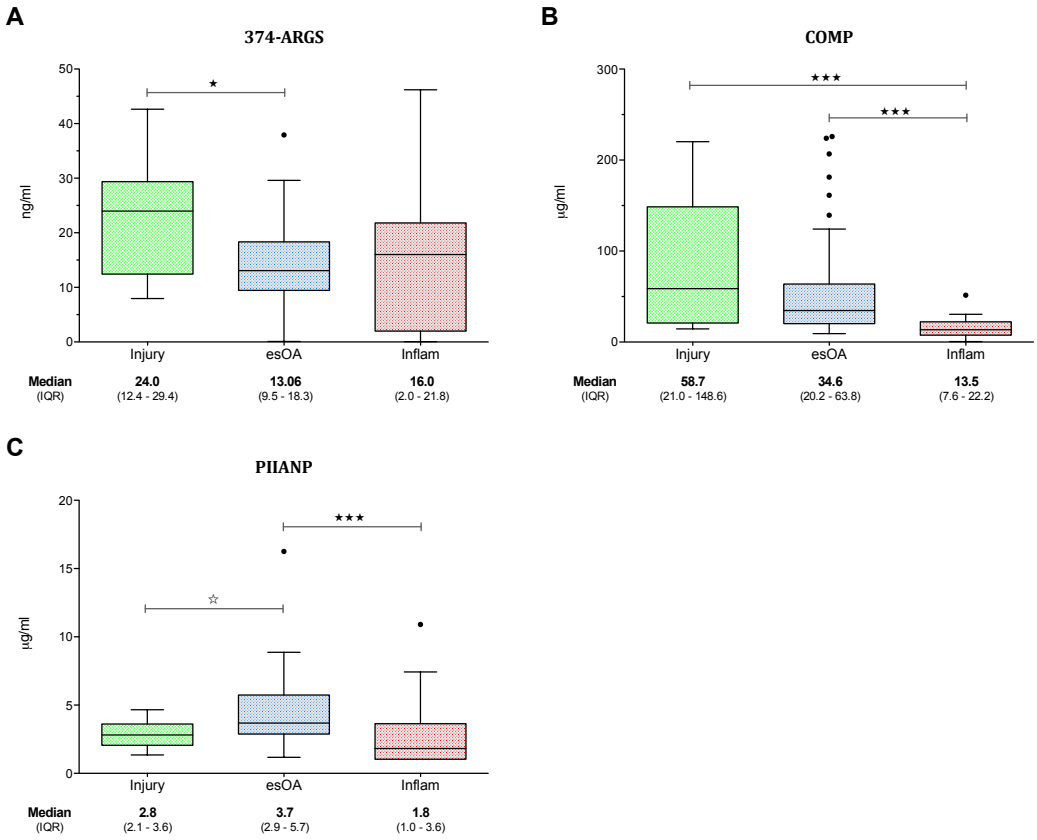


Figure 9.4 – Comparison of markers of cartilage metabolism between cohorts. **[A]** 374-ARGS aggrecan neo-epitope, **[B]** COMP and **[C]** PIIANP. Tukey box-and-whiskers are shown for knee injury (Injury), end-stage knee OA (esOA) and inflammatory (Inflam) cohorts. Measurements outside the Tukey whiskers are shown as black dots. Concentrations are given in µg/ml except 374-ARGS that is given in ng/ml. Statistical significance is denoted by ★ ($p<0.05$), ★★ ($p<0.01$) and ★★★ ($p<0.001$). Clear symbol (i.e. ☆) denotes significance lost after Bonferroni correction for multiple comparisons.

	Injury[A]	esOA[B]	Inflam[C]	p-value	Post-hoc grouping
IL-1β	10.5%	40%	100%	<0.001	A < B < C
IL-2	0%	0%	6.7%	0.150	-
GM-CSF	5.3%	0%	53.3%	<0.001	{A ≈ B} < C
IL-1Ra	0%	31.6%	100%	<0.001	A < B < C
IL-4	0%	1.8%	13.3%	0.083	-
IL-10	0%	0%	46.7%	<0.001	{A ≈ B} < C
IL-2R	0%	33.3%	93.3%	<0.001	A < B < C
RANTES	36.8%	24.6%	80%	<0.001	{A ≈ B} < C
MIP-1α	0%	36.8%	86.7%	<0.001	A < B < C
MIP-1β	0%	5.3%	46.7%	<0.001	{A ≈ B} < C
MIG	0%	10.5%	60%	<0.001	A < B < C
BMP-2	0%	20.7%	55.6%	0.001	A < B < C
BMP-7	0%	0%	0%	-	-
MMP-13	5.0%	46.6%	46.2%	0.003	A < {B ≈ C}
LIGHT	0%	1.7%	83.3%	<0.001	{A ≈ B} < C

Table 9.3 – Proportion of SF samples above LOQ for analytes not qualified for quantitative concentration analysis. The proportion of samples above LOQ is given as a percentage for each cohort: End-stage knee OA (esOA), knee injury and inflammatory arthritis (Inflam). Proportions were compared using the χ-squared test (Monte Carlo method) with post-hoc testing by the Marascuilo procedure. P<0.05 was considered statistically significant. All significant p-values were resilient to adjustment by the Holm step-down method. Significant grouping structure following post-hoc comparisons were A<B<C, {A≈B}<C or A<{B≈C}.

9.4.3 Categorical Markers

The proportions of samples that were quantifiable for IL-1β, regulatory cytokines IL-1RA and IL-2R, chemokines MIP-1α and MIG and the growth factor BMP-2 showed a significant trend being greatest in inflammatory patients, followed by end-stage knee OA patients and least in knee injury patients [Table 9.3]. Inflammatory cytokine GM-CSF, regulatory cytokine IL-10, chemokines RANTES, MIP-1β, and LIGHT were above LOQ in significantly higher proportions of inflammatory patients than end-stage knee OA and knee injury patients, between which there were no significant differences. MMP-13 was quantifiable in significantly fewer knee injury samples than end-stage knee OA and inflammatory samples, which had equivalent proportions. None or almost no samples were above LOQ for IL-2, IL-4 and BMP-7, irrespective of study cohort.

The proportions of samples that were >LOD for the cytokine IL-1RA and chemokines MIP-1α, MIP-1β and MIG was greatest in inflammatory patients, followed by end-stage knee OA patients and least in knee injury patients [Table 9.4]. Inflammatory cytokines

	Injury^{A}	esOA^{B}	Inflam^{C}	p-value	Post-hoc grouping
IL-1β	63.2%	90.0%	100%	0.004	{A ≈ B} < C
IL-2	0%	8.8 %	33.3%	0.008	{A ≈ B} < C
GM-CSF	21.1%	17.5%	66.7%	0.001	{A ≈ B} < C
IL-1Ra	0%	64.9%	100%	<0.001	A < B < C
IL-4	0%	1.8%	93.3%	<0.001	{A ≈ B} < C
IL-10	5.3%	10.5%	93.3%	<0.001	{A ≈ B} < C
IL-2R	26.3%	94.7%	93.3%	<0.001	A < {B ≈ C}
RANTES	42.1%	54.4%	93.3%	0.006	{A ≈ B} < C
MIP-1α	0%	40.4%	100%	<0.001	A < B < C
MIP-1β	5.3%	33.3%	73.3%	<0.001	A < B < C
MIG	5.3%	33.3%	80%	<0.001	A < B < C
BMP-2	41.2%	62.1%	66.7%	0.226	-
BMP-7	11.8%	10.3%	5.6%	0.791	-
MMP-13	5.0%	50.0%	46.2%	0.002	A < {B ≈ C}
LIGHT	17.6%	13.8%	88.9%	<0.001	{A ≈ B} < C

Table 9.4 – Proportion of SF samples above LOD for analytes not qualified for quantitative concentration analysis. The proportion of samples above LOD is given as a percentage for each cohort: End-stage knee OA (esOA), knee injury and inflammatory arthritis (Inflam). Proportions were compared using the χ-squared test (Monte Carlo method) with post-hoc testing by the Marascuilo procedure. P<0.05 was considered statistically significant. All significant p-values were resilient to adjustment by the Holm step-down method. Significant grouping structure following post-hoc comparisons were A<B<C, {A≈B}<C or A<{B≈C}.

IL-1β and IL-2, regulatory cytokines IL-4 and IL-10, RANTES and LIGHT were detectable in significantly higher proportions of inflammatory patients than end-stage knee OA and knee injury patients, between which there were no significant differences. IL-2R and MMP-13 were quantifiable in significantly fewer knee injury samples than end-stage knee OA and inflammatory samples, which had equivalent proportions. There was no difference in detection frequency between cohorts for BMP-2 & BMP-7.

9.4.4 Multivariate Analysis of SF Biology

Multivariate analysis was conducted for 20 quantitative and 12 categorical SF analytes [Table 9.5]. Quantitative markers included 5 inflammatory cytokines (TNF-α, IL-6, IL-8, IL-12 and IL-15); 3 chemokines (MCP-1, IP-10 and Eotaxin), 3 isoforms of TGF-β, 5 matrix enzymes (MMP-1, MMP-3, MMP-9, TIMP-1 and ADAMTS-4); 3 markers of cartilage metabolism (374-ARGS neo-epitope, COMP and PIIANP); and DcR3. Categorical markers included 2 inflammatory cytokine (IL-1β and GMCSF);

	Quantitative	Categorical
Inflammatory cytokines	TNF-α, IL-6, IL-8, IL-12, IL-15	IL-1β, GMCSF
Regulatory cytokines		IL-1Ra, IL-10, IL-2R
Chemokines	MCP-1, IP-10, Eotaxin	RANTES, MIP-1α, MIP-1β, MIG
Growth Factors	TGF-β1, TGF-β2, TGF-β3	BMP-2
Matrix Enzymes	MMP-1, MMP-3, MMP-9, ADAMTS-4, TIMP-1	MMP-13
Cartilage Turnover	374-ARGS, COMP, PIIANP	
Others (bone)	DcR3	LIGHT

Table 9.5 – Quantitative and Categorical markers used for multivariate analysis.

3 regulatory cytokines (IL-1Ra, IL-10 and IL-2R); 4 chemokines (RANTES, MIP-1α, MIP-1β and MIG); growth factor BMP-2, matrix enzyme MMP-13 and LIGHT. Three analytes (IL-2, IL-4 and BMP-7) were excluded because they were quantifiable in less than 25% of patients in any group [Table 9.3].

PCA observation score plot showed evidence of class separation, particularly for quantitative analytes [not shown]. Analysis therefore proceeded to PLS-DA.

Quantitative SF Analytes

PLS-DA using quantitative SF analytes produced good class separation between the 3 study cohorts [Figure 9.5]. A 3-component model was generated that explained 77.0% (R^2=0.770) of the variability between patient groups with a predictive quality of 71.8% (Q^2=0.718). Only two patients with knee injury were misclassified as having end-stage knee OA and all remaining patients were classified correctly. Both these patients had ACL injury plus meniscal tear. The model had excellent accuracy, reliability and specificity [Table 9.6]. All 10 validation test patients were classified correctly as end-stage knee OA (accuracy 100%; reliability 100%).

All quantitative markers were important (VIP>0.8) for the overall model except DcR3, COMP and 374-ARGS that were potentially important (0.8>VIP>0.5), and Eotaxin which was not important (VIP<0.5) [Figure 9.5C]. The loading bi-plot [Figure 9.5B] suggests PIIANP has a strong discriminatory function for end-stage knee OA. TIMP-1,

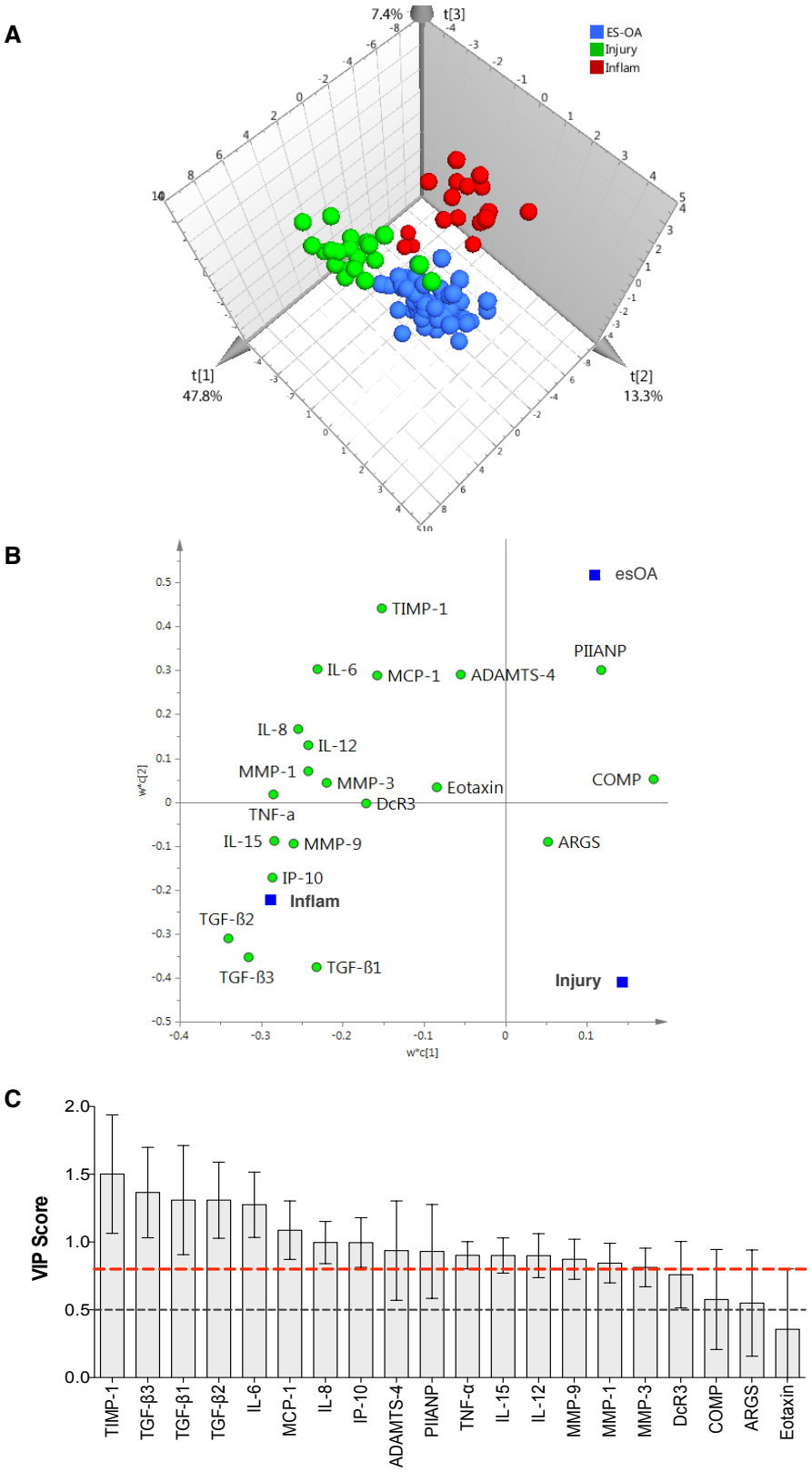


Figure 9.5 – PLS-DA model using 20 quantitative SF analytes. [A] Observation Score Plot. **[B]** Variable Loading Bi-plot. **[C]** Variable Importance for Projection (VIP) Scores. Three-component model: R^2=0.770; Q^2=0.718. esOA - End-stage knee OA; Injury - Knee injury; Inflam - Inflammatory arthritis.

		PREDICTED				
		esOA	**Injury**	**Inflam**	**NC**	**Total$_{Act}$**
ACTUAL	**esOA**	**60**	0	0	0	60
	Injury	2	**18**	0	0	20
	Inflam	0	0	**18**	0	18
	NC	0	0	0	**0**	0
	Total$_{Pred}$	62	18	18	0	

		esOA	Injury	Inflam	Average
(Sensitivity)	**Accuracy**	100%	90%	100%	96.7%
(PPV)	**Reliability**	96.8%	100%	100%	98.9%
	Specificity	94.7%	100%	100%	98.2%

Table 9.6 – Confusion matrix for prediction of end-stage knee OA (esOA), knee injury (Injury) and inflammatory arthritis (Inflam) by quantitative SF analytes. PLS-DA 3-component model (R^2=0.770; Q^2=0.718). Overall Accuracy=98.0%. NC - no classification. PPV - positive predictive value.

ADAMTS-4, MCP-1 and IL-6 also load towards end-stage knee OA. The majority of markers discriminate against knee injury except 374-ARGS that loads in the direction of knee injury patients.

The marker profile for each cohort according to (significant) regression coefficients is shown in [Figure 9.6]. The fingerprint for end-stage knee OA was characterised by ridges of TIMP-1, IL-6, PIIANP, MCP-1, ADAMTS-4 and IL-12 in combination with troughs of TGF-β isoforms, IP-10, IL-15 and MMP-9 [Figure 9.6A].

Streamlined Model

The optimum streamlined PLS-DA model used 8 quantitative SF markers: inflammatory cytokine IL-6; chemokines MCP-1 and IP-10; TGF-β3; aggrecanase ADAMTS-4; metalloproteinase inhibitor TIMP-1 and cartilage metabolism markers PIIANP and COMP. Good class separation was achieved with a 2-component model that explained 69.4% (R^2=0.694) of the variability between patient groups with a predictive quality of 67.3% (Q^2=0.673) [Figure 9.7A]. The streamlined model also had excellent accuracy, reliability and specificity with only 3 knee injury patients being misclassified as end-stage knee OA and all other patients were classified correctly [Table 9.7]. All 3 misclassified injury patients had ACL injury plus meniscal tear. Again all 10 validation

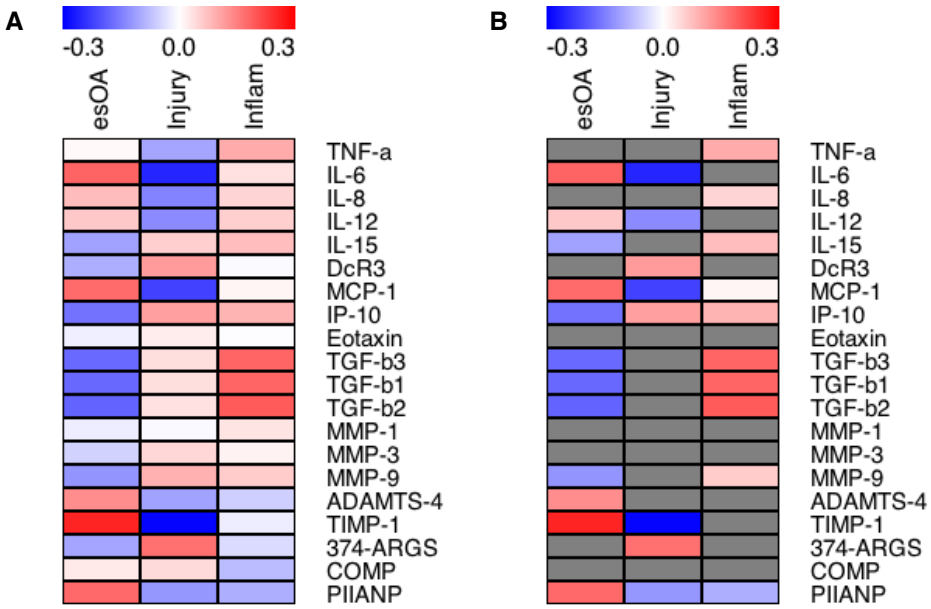


Figure 9.6 – Variable regression coefficient plots for the full PLS-DA model. Regression coefficients for 20 quantitative markers (3-components; R^2=0.770; Q^2=0.718). **[A]** All coefficients **[B]** Significant coefficients only. Grey boxes denote non-significant regression coefficients whose 95% confidence intervals span zero. esOA - End-stage knee OA; Injury - Knee injury; Inflam - Inflammatory arthritis.

test patients were classified correctly as end-stage knee OA (accuracy 100%; reliability 100%).

The loading bi-plot shows PIIANP discriminates best for end-stage knee OA and most markers discriminate against knee injury [Figure 9.7B]. TIMP-1 and ADAMTS-4 also load favourably towards end-stage knee OA. The VIP scores for all 8 markers in the streamlined model were important (TGF-β3, TIMP-1, IL-6, IP-10, MCP-1 and PIIANP) or potentially important (ADAMTS-4 & COMP) [Figure 9.7C]. The fingerprint for end-stage knee OA comprised ridges of PIIANP, TIMP-1, ADAMTS-4 and MCP-1 in combination with troughs of IP-10 and TGF-β3 [Figure 9.8].

Categorical SF Analytes

Categorical analytes produced a 3-component model that explained 54.0% (R^2=0.540) with a predictive quality of 50.8% (Q^2=0.508). The model did not discriminate well between end-stage knee OA and injury, although inflammatory samples showed good separation [Figure 9.9A]. Fourteen knee injury patients were misclassified as end-stage

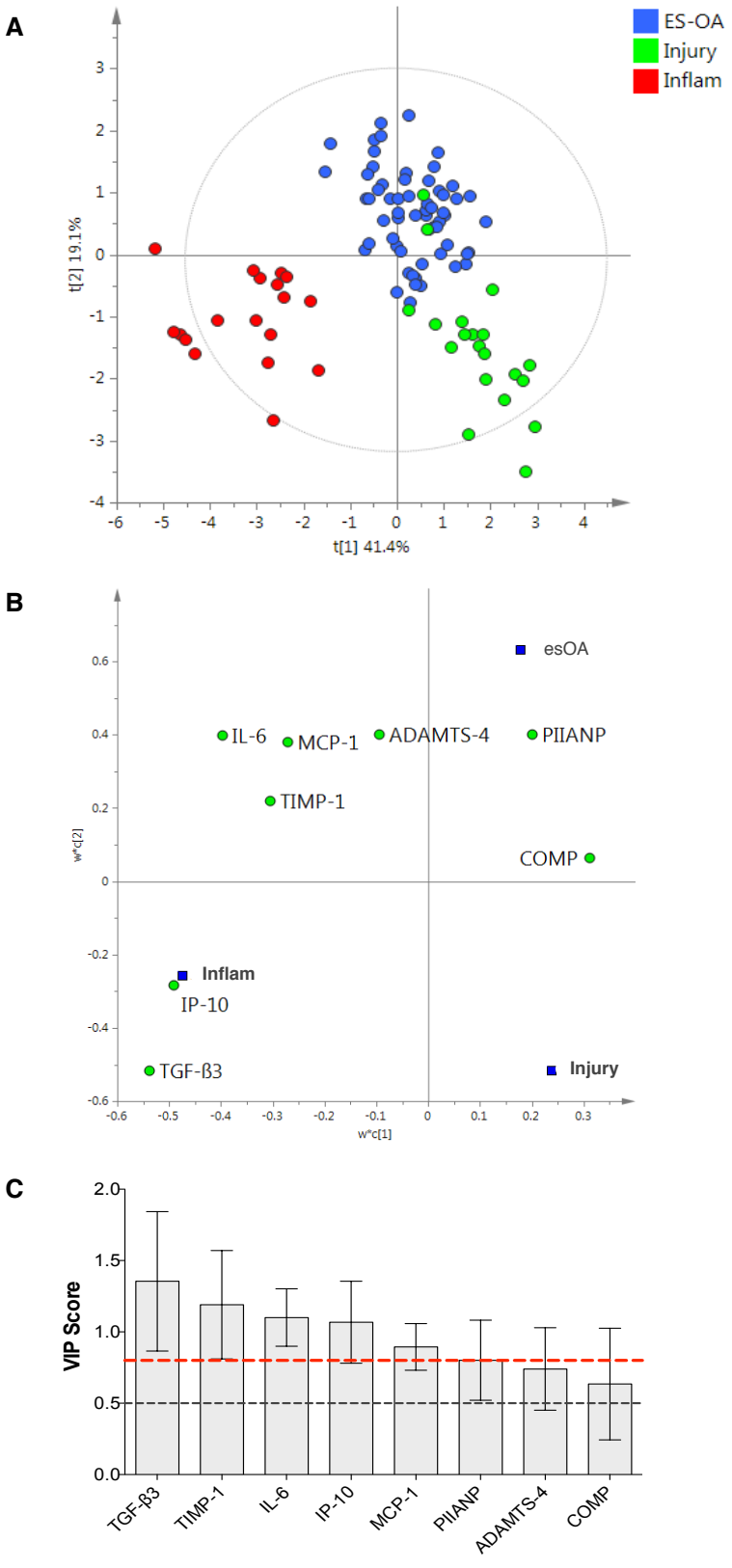


Figure 9.7 – Streamlined PLS-DA model using 8 quantitative SF analytes. **[A]** Observation Score Plot. **[B]** Variable Loading Bi-plot. **[C]** Variable Importance for Projection (VIP) Scores. Two-component model: R^2=0.694; Q^2=0.673. esOA - End-stage knee OA; Injury - Knee injury; Inflam - Inflammatory arthritis.

		PREDICTED				
		esOA	**Injury**	**Inflam**	**NC**	**Total$_{Act}$**
ACTUAL	**esOA**	**60**	0	0	0	60
	Injury	3	**17**	0	0	20
	Inflam	0	0	**18**	0	18
	NC	0	0	0	**0**	0
	Total$_{Pred}$	63	17	18	0	

		esOA	**Injury**	**Inflam**	**Average**
(Sensitivity)	**Accuracy**	100%	85%	100%	95%
(PPV)	**Reliability**	95.2%	100%	100%	98.4%
	Specificity	92.1%	100%	100%	97.4%

Table 9.7 – Confusion matrix for prediction of end-stage knee OA (esOA), knee injury (Injury) and inflammatory arthritis (Inflam) by streamlined quantitative SF analytes. PLS-DA 2-component model (R^2=0.694; Q^2=0.673). Overall Accuracy=96.9%. NC - no classification. PPV - positive predictive value.

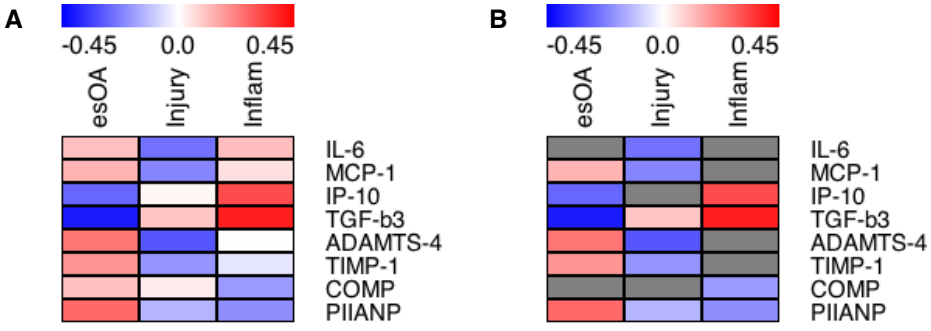


Figure 9.8 – Variable coefficient plots for streamlined PLS-DA model. Streamlined PLS-DA model coefficients for 8 quantitative markers (2-components; R^2=0.694; Q^2=0.673). **[A]** All coefficients **[B]** Significant coefficients only**.** Grey boxes denote non-significant regression coefficients whose 95% confidence intervals span zero. esOA - End-stage knee OA; Injury - Knee injury; Inflam - Inflammatory arthritis.

		PREDICTED				
		esOA	**Injury**	**Inflam**	**NC**	**Total$_{Act}$**
ACTUAL	**esOA**	**58**	2	0	0	60
	Injury	14	**6**	0	0	20
	Inflam	2	0	**16**	0	18
	NC	0	0	0	**0**	0
	Total$_{Pred}$	74	8	16	0	

		esOA	**Injury**	**Inflam**	**Average**
(Sensitivity)	**Accuracy**	96.7%	30%	88.9%	71.9%
(PPV)	**Reliability**	78.4%	75%	100%	84.5%
	Specificity	57.9%	97.4%	100%	85.1%

Table 9.8 – Confusion matrix for prediction of end-stage knee OA (esOA), knee injury (Injury) and inflammatory arthritis (Inflam) by categorical SF analytes. PLS-DA 3-component model (R^2=0.540; Q^2=0.508). Overall Accuracy=81.6%. NC - no classification. PPV - positive predictive value.

knee OA leaving only 6 correctly classified [Table 9.8]. Model performance was less satisfactory, particularly the accuracy for classifying injury and specificity for end-stage knee OA.

The majority of categorical SF markers loaded towards inflammatory patient group [Figure 9.9B]. All markers were important (VIP>0.8) for the model except for MIP-1β and BMP-2, which were potentially important (0.8>VIP>0.5) [Figure 9.9C]. The coefficient plots confirm that most of the categorical markers discriminate in favour of inflammatory samples [Figure 9.10]. However, a combination of quantifiable MMP-13 and IL-1RA with unquantifiable LIGHT, RANTES, GMCSF, IL-10 and MIG may help characterise end-stage knee OA.

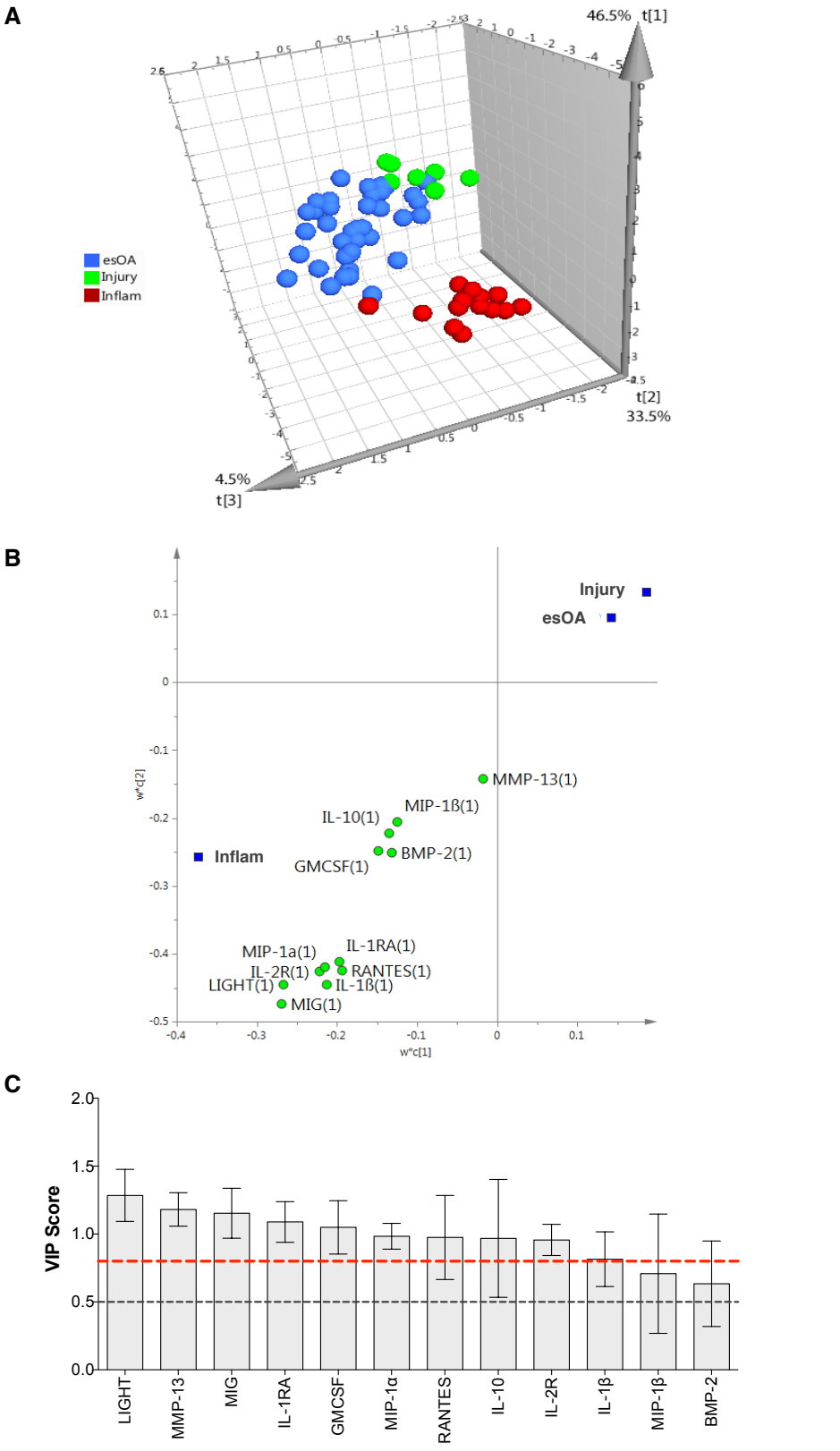


Figure 9.9 – PLS-DA model using 12 categorical SF analytes. [A] Observation Score Plot. **[B]** Variable Loading Bi-plot. **[C]** Variable Importance for Projection (VIP) Scores. Three-component model: R^2=0.540; Q^2=0.508. esOA - End-stage knee OA; Injury - Knee injury; Inflam - Inflammatory arthritis.

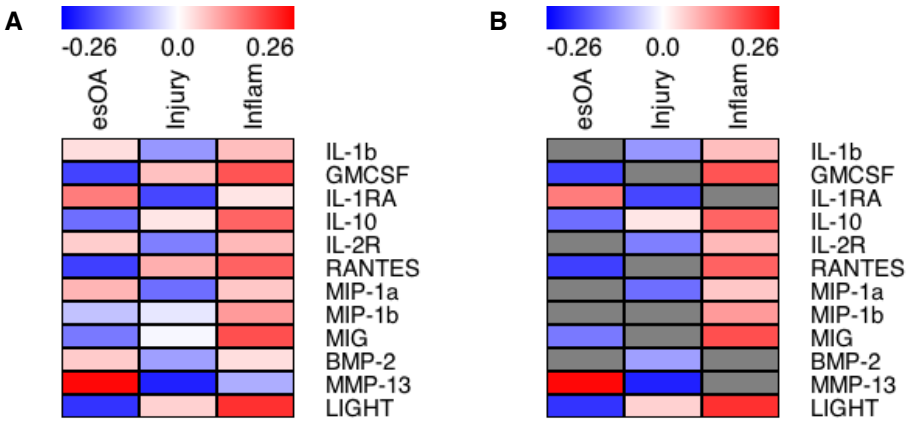


Figure 9.10 – Variable coefficient plots for categorical analyte PLS-DA model. PLS-DA model coefficients for 12 categorical markers (3-components; R^2=0.540; Q^2=0.508). **[A]** All coefficients **[B]** Significant coefficients only**.** Grey boxes denote non-significant regression coefficients whose 95% confidence intervals span zero. esOA - End-stage knee OA; Injury - Knee injury; Inflam - Inflammatory arthritis.

9.5 Discussion

This study aimed to discover a SF biomarker signature or fingerprint for end-stage knee OA from a wide spectrum of physiologically relevant cytokines, chemokines, growth factors, matrix enzymes and markers of cartilage turnover. Single marker comparisons between study cohorts suggested elevated PIIANP and reduced TGF-β may be potential biomarkers for end-stage knee OA. Multivariate analysis produced and (partially) validated both detailed and streamlined models with cohort-specific fingerprints that accurately and reliably discriminated end-stage knee OA from knee injury and inflammatory arthritis with almost 100% efficacy.

Between-group comparisons for individual markers mostly followed a stepwise concentration pattern of knee injury < end-stage knee OA < inflammatory arthritis. Although this provides internal validation for SF measurements, they are largely uninformative in the absence of cohort-specific reference ranges. The presentation of reference ranges has been intentionally avoided in this study because measured concentrations are at best quasi-quantitative being dependent on the precise assay used and numerous pre- and post analytical factors (Chapters 3 and 4). PIIANP, TGF-β1 and TGF-β3 did show some evidence of differential SF expression in end-stage OA. These markers will be discussed later.

This is the first study to combine wide-spectrum immunoassay examination of SF with PLS-DA multivariate "-omics" statistical analysis in OA biomarker research. This novel hybrid approach produced an initial full model from 20 quantitative analytes with R^2(goodness-of-fit) and Q^2(predictive quality) parameters >0.7, which is excellent for a biological system. This full model generated an end-stage knee OA fingerprint comprising "ridges" of TIMP-1, IL-6, PIIANP, MCP-1, ADAMTS-4 and IL-12 with "troughs" of TGF-β isoforms, IP-10, IL-15 and MMP-9. A streamlined model with 8 quantitative analytes had R^2(goodness-of-fit) and Q^2(predictive quality) parameters of ~0.7, and produced a simplified fingerprint comprising "ridges" of PIIANP, TIMP-1, ADAMTS-4 and MCP-1 with "troughs" of IP-10 and TGF-β3. Both models achieved excellent (>95%) diagnostic accuracy, reliability and specificity, and were successfully validated with new end-stage knee OA test samples. In contrast, the PLS-DA model derived from 12 categorical analytes that did not qualify for quantitative analysis, was not able to distinguish well between end-stage knee OA and knee injury patients, but discriminated well for inflammatory patients. Both R^2(goodness-of-fit) and Q^2(predictive quality) parameters were only ~0.5 and therefore further interpretation of this model must be done with caution.

It must be highlighted that although the focus of this thesis is OA, each model includes parallel fingerprints for knee injury and inflammatory arthritis. By this virtue the multivariate models have equivalent diagnostic efficacy for these conditions. The rationale for including patients with knee injury and inflammatory arthritis was to provide internal (construct) validity to SF measurements in the absence of normal control samples, and to provide an interesting biological contrast to OA from other knee joint pathologies. With a few exceptions, most OA biomarker studies use a single non-OA control cohort. This is usually "normal" patients in studies using peripheral biofluids, or patients with knee injury in SF studies. Differences in markers of matrix (cartilage or bone) turnover, matrix enzymes and soluble mediators such as inflammatory cytokines, growth factors etc. may not be specific OA biomarkers because they may be differentially expressed in other joint pathologies and in particular inflammatory arthritis. One only needs to peruse the inflammatory arthritis literature where many studies use OA as the reference control cohort for biomarker discovery. Therefore many putative OA biomarkers may be indicators of synovitis and inflammation common to OA and

inflammatory arthritis. The inclusion of 2 non-OA comparator groups, one on either side of the OA disease spectrum, significantly increased the specificity of potential biomarkers and boosted the statistical power of analysis of relatively few (~100) patients permitting 100% accuracy and 95% reliability of end-stage OA diagnosis with only 8 markers.

The observation score plots of the PLS-DA models demonstrated far separation of the inflammatory group from both knee injury and end-stage OA groups, between which class separation was distinct but less clean. This was particularly so for the model using categorical analytes, but may be expected since these comprised mostly cytokines and chemokines. In the higher quality full and streamlined multivariate models, using a more diverse range of quantitative analytes, 2 and 3 patients with knee injury, respectively, were misclassified as end-stage knee OA. Both misclassified patients in the full model were also misclassified in the streamlined model in addition to a third patient. Although by no means normal, patients with knee injuries had no radiological (radiographic and MRI) and arthroscopic evidence of OA. However, this was a heterogeneous group including isolated meniscal injury and ACL injury +/- meniscal injury and it is worth pointing out that all 3 misclassified patients had ACL injury and meniscal tear. The relative proximity of knee injury patients to those with end-stage knee OA suggests that these groups may lie on a biological continuum. A persistent abnormal SF biology associated with molecular and pre-macroscopic OA changes may render injury patients as a high-risk/pre-OA cohort. There is strong evidence that such injuries are a very important risk factor for OA and Lohmander has described these individuals as “the young patient with an old knee” [reviewed in 24, 25]. The development of OA is 10 times greater in the ACL injured patient than normal individuals regardless of treatment. At 10–12 years following diagnosis, an average of 50% of patients diagnosed with an ACL or meniscus tear develop OA with associated pain and functional impairment regardless of surgical intervention. The exact reason for this is still uncertain, and may be due to meniscal, articular cartilage or subchondral bone damage at the time of injury. Furthermore, many animal models of OA involve cruciate transection or meniscal damage. Another possibility raised by the relative proximity of the injury and end-stage knee OA cohorts is that the multivariate models and OA fingerprint may have extended utility for earlier stages of OA and not

just end-stage OA. However, this requires formal testing of the current models with appropriate patient samples.

Although the quantitative PLS-DA model and their fingerprints had high diagnostic accuracy and reliability, external validation with new test samples is essential. Partial validation was achieved with 10 patients with end-stage knee OA (lateral compartment OA), of which all were correctly classified (100% accuracy, 100% reliability). However, full validation requires further knee injury and inflammatory SF samples. Although, Chapter 8 demonstrated no difference in the SF biology of different structural phenotypes of end-stage knee OA by standard and MVDA methods, further validation of the multivariate models with more phenotype-specific end-stage knee OA samples would be reassuring. The model derived from categorical analytes was less convincing and highlights the need for more sensitive immunoassays to test the biological importance of these markers. The multivariate models and biological fingerprints described in this study are by no means definitive and future work to refine the models in terms of the number and nature of biological mediators is actively encouraged.

This study represents a potential milestone in OA biomarker discovery. Therefore, the biomarker fingerprints need biological qualification and one needs to discuss the multivariate approach and potential implications of the findings.

Biological Qualification

It must be emphasised that the multivariate model fingerprints describes a pattern of biological features in combination with each other and must be taken as a whole. However, biological qualification of the findings requires examination of the individual markers.

Cytokines and Inflammatory Markers

As expected, pro-inflammatory and regulatory cytokines largely loaded towards inflammatory arthritis in both quantitative and categorical multivariate models [414, 415]. This was echoed in the single-marker comparisons, where concentrations were greatest

for inflammatory samples, followed by end-stage knee OA and then non-OA knee injury. Similar findings have been reported previously [416, 417]. Although there is substantial evidence for the role of inflammatory mediators and synovitis in the pathophysiology of OA, SF cytokines measurements in the multivariate models were more specific for inflammatory arthritis and important for discriminating against non-OA knee injury. Consequently, they may not be ideal OA biomarkers. Their individual utility would require reference ranges describing OA; below which OA is unlikely and above which inflammatory disease is more likely. The presentation of reference ranges has been intentionally avoided in this study because measured concentrations are at best quasi-quantitative being dependent on the precise assay used and numerous pre- and post analytical factors (Chapters 3 and 4). The diminished relative merit of cytokines as specific OA biomarkers is further suggested by the finding that only IL-6 was retained in the streamlined PLS-DA model with reduced levels characterising knee injury. IL-6 may have a specific role in OA by inhibiting type II collagen gene expression (COL2A1) by chondrocytes [418]. However, the inflammatory cytokine IL-15 was a significant negative marker in the end-stage knee OA fingerprint in the full quantitative PLS-DA model. In addition, the difference in SF IL-15 between knee injury and end-stage knee OA was not significant after correction for multiple comparisons. These findings are consistent with a study by Scanzello et al. that demonstrated significantly lower SF IL-15 levels in end-stage knee OA compared to early knee OA (degenerate meniscal tears and medial KLG <2) [419]. IL-15 is predominantly a membrane bound mediator and increased intra-articular detection of the soluble form in early OA may potentially represent reduced biological activity.

Eight of the 15 analytes that did not qualify for quantitative analysis were cytokines: 4 inflammatory cytokines IL-1β, IL-2, GMCSF and LIGHT, and nearly all regulatory cytokines on the biological panel (IL-1RA, IL-2R IL-4 & IL-10). The quantification and detection frequencies largely followed the anticipated pattern demonstrated by the quantitative cytokines. As expected, the categorical cytokines included in the multivariate model were mostly associated with inflammatory arthritis and either opposed or were not significant for end-stage knee OA and knee injury. A notable exception was IL-1RA, which was a significant positive element in the OA fingerprint. This is consistent with a study by Richette et al reporting higher IL-1RA/IL-1β SF

concentration ratios in patients with knee OA than rheumatoid arthritis [420]. Both IL-1β and IL-1RA were negative elements in the knee injury fingerprint in keeping with the catabolic role of IL-1β implicated in the initiation and progression of OA [421]. Curiously, both LIGHT and its decoy receptor DcR3 were positive elements in the knee injury fingerprints even though LIGHT was not quantifiable in any knee injury patients and DcR3 concentrations were equivalent to end-stage knee OA. It is difficult to offer a biological explanation for these findings since both markers are usually associated with inflammatory arthritis [422, 423, 424, 425]. The regression coefficient for LIGHT was very low (<0.05) and that for DcR3 just achieved significance (95%CI 0.22 – 0.01) and therefore they may instead be artefacts of the multivariate models that further sample analysis would eliminate. Although one must be cautious in over-interpretation of categorical data, it is reassuring that IL-10 is a negative element in the end-stage knee OA fingerprint since this regulatory cytokine has been shown to have chondroprotective properties [426].

Chemokines

Chemokines play a crucial role in inflammation, but their actions are complex and not just restricted to recruitment of inflammatory cells. Accordingly, quantitative and categorical chemokine measurements were, on the whole, greatest in inflammatory samples. This pattern has been demonstrated previously [427]. MCP-1 and IP-10 were important markers in both the full and streamlined models being significant positive and negative elements, respectively, in the OA fingerprints. A recent comprehensive study by Harris et al. demonstrated that elevated SF MCP-1 levels in OA reduces the chondrogenic potential of synovial mesenchymal progenitor cells at a gene and protein level [428]. A preliminary extensive multiplex analysis of SF by the authors showed significantly higher levels of MCP-1 in OA patients than cadaveric controls. The authors suggest that prolonged exposure to elevated MCP-1 “locks the joint in a vicious cycle of ineffective repair”. In another wide-ranging multiplex analysis of SF from cadaveric normal controls and patients with mild/moderate OA or severe OA, Heard et al. found MCP-1 to be the most influential cytokine discriminating between normal and OA samples [429]. Saetan et al. demonstrated that both SF

and serum IP-10 were inversely correlated with increasing radiographic severity in patients with knee OA and therefore offers an explanation for this chemokine being an important negative feature in the end-stage knee OA fingerprint [430]. IP-10 is also a strong positive element in the parallel inflammatory fingerprint in both the full and streamlined multivariate models. This is reflected by measured SF concentrations of IP-10 being 2 orders of magnitude greater in inflammatory samples than end-stage knee OA and knee injury between which there was no difference. Elevated IP-10 is known to be a feature of rheumatoid and psoriatic arthritis [431, 432, 433]. There was no difference in eotaxin(-1) between study cohorts and consequently this chemokine was not important in the multivariate models (VIP$<$0.5) and did not feature in the biological fingerprints. Eotaxin has been induced in cytokine-activated chondrocytes and can stimulate chondrocyte MMP-3 expression and release [434, 435]. However, there are conflicting reports in the literature with both elevated and unquantifiable SF eotaxin in OA patients compared to controls [428, 241]. It is possible that increased clearance rates associated with greater synovitis from injury to OA to inflammatory knees masked any measured differences or the assay itself may have been unreliable. The categorical chemokines included in the multivariate models were mostly associated with inflammatory arthritis and either opposed or were not significant for end-stage knee OA and knee injury. The detection and quantification frequencies of RANTES were not significantly different between end-stage OA and knee injury patients, which is consistent with a similar recent comparison [436].

Growth Factors

Osteophytosis is a cardinal feature of OA and TGF-β has been regarded as central to osteophyte formation [reviewed in 56]. At first glance, it is surprising that all three TGF-β isoforms, and in particular TGF-β3, were negative elements of the end-stage knee OA fingerprint. SF measurements of TGF-β1 and TGF-β3 were lower in end-stage knee OA than knee injury and there was no difference for TGF-β2. These results may be a consequence of loss of cartilage volume in end-stage knee OA since a major source of TGF-β resides in the cartilage matrix [437, 438]. Immunohistochemical analysis of OA cartilage in animal models have also demonstrated matrix TGF-β depletion [439].

However, the results of Chapter 8 do not show any differences in SF TGF-β (or any other marker) levels in cumulatively more severe structural phenotypes of end-stage knee OA. Recent work has illustrated the split-personality of TGF-β that has complex pleiotropic actions in OA including both pro-inflammatory/catabolic and regulatory/anabolic capabilities depending on differential receptor pathway signalling [reviewed in 440, 53, 441, 57]. TGF-β is required to maintain normal cartilage homeostasis and loss of signalling can lead to degeneration. In particular, TGF-β can regulate MMP and TIMP gene expression via Smad and AP-1 signalling pathways and therefore reduced levels in OA may promote in MMP-mediated matrix degradation [442]. The results of this study favour the paradigm that OA is associated with reduced TGF-β anti-catabolic and reparative function. It is also possible that reduced matrix TGF-β in end-stage knee OA is compounded by increased SF clearance by associated synovitis not found in knee injury. However, all three TGF-β isoforms were significantly higher in inflammatory patients and consequently load towards inflammatory arthritis in the multivariate models. Synovitis, a cardinal feature of inflammatory joint disease, is likely to be a source of TGF-β, which has been shown to be elevated in the SF and synovium of inflammatory arthritis [443, 416, 414].

BMPs have both anabolic and catabolic properties and play a role in osteophyte formation, cartilage repair and remodelling [reviewed in 444, 54]. Unfortunately, BMP-2 and BMP-7 did not qualify for quantitative analysis with BMP-7 virtually undetectable in SF samples from all study cohorts. However, other studies have successfully measured BMP-7 in OA and inflammatory SF using immunoassays [445, 446, 417] and therefore the custom assay used in this study may have unsatisfactory analytical sensitivities. This again highlights the danger in directly comparing studies using different assays. Quantification frequency of BMP-2 was higher in inflammatory, followed by end-stage knee OA and then knee injury samples, where no samples were >LOQ. Higher quantification frequency in inflammatory samples is most likely the result of BMP-2 expression by inflammatory synovium and stimulation by pro-inflammatory cytokines [447, 448, 449]. BMP-2 was a negative element in the knee injury fingerprint for categorical markers, but did not feature in the parallel end-stage knee OA or inflammatory fingerprints. These results are consistent with reduced BMP-2 mRNA and protein expression in healthy versus OA cartilage [450, 447, 439].

Matrix Enzymes

Matrix enzymes such as MMPs, aggrecanases and TIMPs are key players on the frontline of the battle between matrix degradation and synthesis in OA where disequilibrium ultimately favours matrix destruction. Of the assayed MMPs, it was surprising that only MMP-9 (gelatinase-B) and MMP-13 (collagenase-3) featured in the multivariate models. The importance of gelatinases in OA is unclear [451]. MMP-9 was an important variable in the full multivariate model, where it was a negative element in the OA fingerprint and a positive element in the parallel inflammatory fingerprint, but was not present in the streamlined model. Furthermore, SF MMP-9 concentrations were significantly higher in inflammatory samples than end-stage knee OA and knee injury, which were equivalent. The literature largely supports the above findings with several studies demonstrating higher SF MMP-9, as well as MMP-1 (collagenase-1), MMP-3 (stromolysin-1) and TIMPs, in inflammatory arthritis than OA [452, 453, 454, 455, 456]. A recent study by Ryu et al. demonstrated significantly greater MMP-9 activity in SF from inflammatory knee conditions than OA using near infrared fluorescence probes and gelatin zymography [457]. MMP-13 (collagenase 3) has a potent ability to degrade type II collagen and therefore its pivotal role in OA pathophysiology and potential as a therapeutic target have received considerable attention [43]. It is therefore not surprising that MMP-13 featured as a positive element in the OA fingerprint and a negative element in the parallel knee injury fingerprint in the multivariate model for categorical markers. MMP-13 did not qualify for quantitative analysis, being quantifiable in nearly 50% of end-stage knee OA and inflammatory samples and only 5% of injury samples. Marini et al. also observed greater MMP-13 activity in patients with advanced cartilage degeneration compared to normal and early changes assessed by arthroscopy [458]. In the absence of sufficient quantitative data, it is difficult to comment on differences in SF MMP-13 levels between OA and inflammatory arthritis and there are conflicting reports in the literature. Yoshihara et al. detected MMP-13 in significantly more knee SF samples from patients with rheumatoid arthritis than isolated OA using ELISA [453]. However, Heard et al. demonstrated higher levels, with ELISA and Luminex, in knee SF of OA than rheumatoid arthritis patients on a variety of medications [456]. One can speculate that MMP-13 would be an important positive element in the OA fingerprint

of the quantitative multivariate models if a more sensitive assay was used.

Neither MMP-1 (collagenase-1) nor MMP-3 (stromolysin-1) was a significant component in any of the biological fingerprints. Although measured SF concentrations of both were greatest in inflammatory samples followed by OA and then knee injury, the utility of these analytes as potential individual biomarkers is compromised by the need for disease-specific reference ranges as discussed above. It must be stressed that immunoassays measure total MMP reflecting a combination of latent pro-MMP, active MMP and both forms complexed to TIMPs and potentially other inhibitors. This study did not measure enzyme activity and, therefore, one cannot dismiss MMP-1 or MMP-3 in one of these specific forms as potential biomarkers.

ADAMTS-4 (aggrecanase-1) was a positive element in the end-stage knee OA fingerprints of both the full and streamlined multivariate models. In the streamlined model it was also a negative element in the knee injury fingerprint. In addition to the finding that MMP-3 did not feature in the OA fingerprints, these results are consistent with the current consensus that the aggrecanases have a more specific and potent role in OA aggrecanase degradation than MMPs [459, 460, 461, 462, 463]. The expression of matrix enzyme inhibitors, such as TIMPs, is known to be elevated in OA as an attempt at maintaining matrix homeostasis. It was therefore not surprising that SF TIMP-1 concentrations mirrored the general pattern for MMPs, and TIMP-1 featured prominently in both the full and streamlined multivariate models being strongly associated with end-stage knee OA and opposing knee injury. The TIMP-1 immunoassay measures free and complexed TIMP-1 and SF TIMP-1 levels may represents a global approximation of matrix enzyme levels. Furthermore, there is increasing evidence that TIMPs, and in particular TIMP-1, have MMP-independent cytokine-like signalling activity involved in various processes including cell growth, apoptosis and differentiation [reviewed in 464].

Matrix Metabolism

In both the full and streamlined multivariate models, PIIANP was a positive element in the end-stage knee OA fingerprint and a negative element in the parallel knee injury

and inflammatory fingerprints. This was reinforced by greater SF PIIANP concentrations in end-stage knee OA patients than the other cohorts, between which there was no difference. PIIANP reflects synthesis of an embryonic form of type II collagen by chondrocytes that have undergone a phenotypic shift in OA cartilage [465, 466]. At the time of writing, there were no previous studies reporting SF PIIANP levels in OA. Serum PIIANP concentrations have been reported to be significantly lower in patients with knee OA and rheumatoid arthritis than healthy controls, and there are conflicting reports regarding the association of serum levels with disease progression [467, 468, 469, 470]. The findings in this study are more consistent with the expected paradigm and, furthermore, suggest SF PIIANP as a potential specific OA biomarker in its own right.

Both aggrecanase-generated aggrecan fragments and COMP release from degenerate cartilage matrix are considered potential biomarkers for OA. However, neither 374-ARGS aggrecan neo-epitope or COMP was a significant element in the detailed OA fingerprint. ARGS was intriguingly a positive element in the parallel knee injury fingerprint and this was echoed in the finding that SF ARGS concentrations were significantly greater in knee injury than end-stage OA samples. SF COMP levels were lowest in inflammatory samples and there was a trend to be greatest in knee injury samples. Although COMP qualified for the streamlined multivariate model, it was only significant as a negative element for the inflammatory fingerprint. These findings are unexpected since SF MMP and ADAMTS-4 concentrations were lowest in knee injury samples. Furthermore, the results question the true biological nature of macroscopically normal cartilage in knee injury patients. The mechanical and biological implications of knee injury for the risk and progression to OA are well documented in the literature. Swärd et al demonstrated elevated SF 393-ARGS aggrecan neo-epitope levels in acutely injured knees up to 23 days after injury in comparison to healthy reference knees [471]. Larsson et al. reported SF 393-ARGS aggrecan neo-epitope levels in injured knees (ACL rupture +/- meniscal tear) with low OA scores were higher than in patients with knee OA within 12 weeks from injury and were equivalent after 12 weeks from injury [472]. However, in the present study SF 374-ARGS levels remains higher in the knee injury than end-stage knee OA patients at a median 6.5 months (IQR 4 – 9.75 months) post index injury. The authors also suggest that elevated SF ARGS is

due to aggrecanase activity against existing matrix aggrecan. In severe and end-stage knee OA, there is reduced existing cartilage matrix volume and therefore relatively lower ARGS neo-epitope release may be expected despite higher aggrecanase levels. This is supported by evidence that SF proteoglycan fragment concentration decreases with increasing radiographic OA severity which is associated with lower cartilage and proteoglycan mass [473].

The findings for COMP are more difficult to reconcile. Kokebie et al. were unable to demonstrate any differences in SF COMP concentration between patients with knee OA, rheumatoid arthritis or asymptomatic donors [417]. El-Arman et al. reported higher SF COMP concentrations in patients with knee OA than traumatic effusion and furthermore there was a positive correlation with both radiographic and MRI severity scores [474]. The traumatic effusion group may represent a different population to the present knee injury cohort since they had no evidence of meniscal injury and the authors make no comment on possible ACL injury. Consistent with the results of the present study, Lohmander et al. demonstrated SF COMP to be higher in knees with meniscal injury or ACL rupture +/- meniscal injury than primary OA [475]. The authors also report a trend for SF COMP levels to decrease with increasing knee OA severity. As with ARGS aggrecan neo-epitope, loss of cartilage matrix volume in severe and end-stage knee OA means there is less COMP for release. COMP that has already been released may have been degraded by MMPs and aggrecanase and therefore not detected by the immunoassay [476, 477, 478, 479, 480]. The negative association of SF COMP with the inflammatory cohort may be also be a consequence of group heterogeneity and disease modifying treatments [481, 482]. Furthermore, increased molecular clearance associated with greater synovitis in end-stage OA and inflammatory arthritis may contribute to lower measured levels of both ARGS aggrecan neo-epitope and COMP compared to knee injury.

Multivariate Approach

This is not the first study to conduct wide-spectrum analysis of a biological fluid for OA biomarker discovery. There is a mounting interest in combined multi-marker approaches due to the failure in finding single biomarkers for a disease as complex as

OA. This has been fuelled, in part, by the rise in multiplex immunoassay platforms and increasing used of proteomic methods. The vast majority of these studies, however, sample peripheral blood or urine. Although SF investigations are relatively few in number, the biomarker community is showing a slow return to focus on SF. However, many of these studies fall victim to at least one of the following limitations: small sample numbers; poor patient characterisation in terms of clinical and/or structural severity; lack of "normal" control samples; few patient cohorts leading to lack of OA specificity; and reversion to multiple single-marker statistical comparisons between study groups often without adjustment for family-wise error (FWE). Further discussion will concentrate on SF studies relevant to the current work.

As illustrated by this work, the use of multivariate data analysis (MVDA) can significantly enhance the usefulness and interpretation of wide-spectrum sample analysis. In this regard some studies represent real missed opportunities, which their authors could potentially rectify by reanalysing the data; in particular, the studies by Huebner & Kraus [483], Kokebie et al. [417] and Rubenhagen et al. [484]. Although an animal study, Huebner & Kraus conducted Luminex analysis of a wide range of 18 mediators in serum and SF from OA prone (Hartley) and OA resistant guinea pigs. However, the authors conducted standard group comparisons with no adjustment for FWE. Kokebie et al. used ELISA to measure a range of cytokines, COMP and bone markers in SF from patients with OA (n=45), RA (n=22) and age–matched asymptomatic cadaveric donors (n=20). OA patients underwent clinical (WOMAC) and radiographic (Kellgren & Lawrence) scoring. The authors conducted only multiple group comparisons and correlations, but must be commended for the use of three patient cohorts for reasons previously discussed. Rubenhagen et al. used Luminex assays to measure a wide range of 16 cytokines, chemokines, growth factors and MMPs in SF from patients with radiographic OA (n=58) and no radiographic OA (n=24). Patients were well stratified according to age and Kellgren & Lawrence grade. The focus of the analysis, however, was the correlations of markers with age and radiographic severity.

The use of MVDA approaches in OA biomarker studies, and in particular principal components analysis (PCA), is increasing irrespective of biological fluid analysed [485, 486, 239, 487, 240, 241, 429]. Other pattern recognition methods that have

been used in SF analysis, albeit in inflammatory arthritis, include random forest and cluster analysis [488, 268, 244]. PCA can be used as an unsupervised classification approach, but is best-suited and most commonly employed for dimension (variable) reduction and grouping of mediators into clusters with potential pathophysiological relevance. Beekhuizen et al recently conducted wide-spectrum screening of 47 markers in patients with end-stage knee OA (n=18) and cadaveric controls (n=16) using several multiplex Luminex assays [241]. Although multiple group comparisons were conducted without correction for FWE, the authors performed PCA to reveal 6 clusters of inter-related mediators, which may have pathophysiological importance. A recent study by Heard et al. measured an extensive range of 42 mediators in SF of patients with severe OA (n=20), mild/moderate OA (n=12) and normal cadaveric donors (n=34) using the Luminex platform [429]. Severe OA was classified radiographically and mild/moderate OA by arthroscopic grading in the absence of radiographic features. An identical serum analysis was conducted in larger patient numbers. The authors conducted multiple pairwise cohort comparisons for each marker without adjustment for FWE and then PCA on significant markers for dimension reduction and repeat cohort comparisons. Unsupervised visual grouping of patients by PCA suggested relatively good separation of severe OA from mild/moderate and normal between which there was notable overlap, particularly in serum. The study by Heard et al. is of particular interest as it one of very few investigations that attempts to use wide-spectrum sample analysis to formally classify patients. Although this is still an unsupervised approach, the authors use K-means clustering on the PCA results to generate 3 groups, presumably representing the original study cohorts. There was good accuracy for serum analysis with 95% normal, 93% mild/moderate OA and 82% severe OA being correctly grouped. However, the performance in SF was poor with 50% normal, 100% mild/moderate OA and 53% severe OA being correctly grouped. In particular, 50% of normal patients were classified as mild/moderate OA and 35% of patients with severe OA were classified as normal. Barring a results table, the authors make no comment on the SF results in their manuscript. The application of the current study's multivariate approach to the immunoassay data from the reports by Beekhuizen et al. and Heard et al. would be highly informative.

Proteomic technologies are increasingly being used in OA biomarker discovery as they

allow extreme high-throughput sample analysis with the promise of combination and potentially novel biomarkers. Multivariate statistical approaches are often imbedded in the computation methods required to interpret proteomic analysis. Proteomic methods applied to SF analysis have been able to demonstrate functionally different protein profiles in OA, RA and healthy patients [269, 489]. Pan et al. analysed SF from patients with OA (n=30, stratified according to radiographic severity) and healthy controls (n=10). There were 61 differential protein peaks between patient groups, of which 2 had high diagnostic value by receiver operating characteristic (ROC) analysis. Furthermore, the authors used a complex multivariate statistical technique, genetic algorithm class prediction, which enabled 100% recognition capacity with 87% internal cross-validation. Although this represents a diagnostic fingerprint, the proteins peaks were not identified. Gobezie et al. used proteomic methods to identify differential protein profiles in SF from early OA confirmed at arthroscopy for meniscal tears (n=21), end-stage knee OA at arthroplasty (n=18) and healthy donors (n=20) [490]. The authors must be commended on their attempts and success at obtaining healthy control samples meeting stringent exclusion criteria without lavage. PCA was performed on 117 identified peptides and revealed separation of healthy from OA patients, but poor distinction between early and end-stage disease. This finding has particular relevance to the present study, suggesting that the end-stage knee OA fingerprint may be of diagnostic value in earlier stages of disease. However, this proposition requires formal future investigation. The authors conducted multiple comparisons, with appropriate adjustment for FWE, between healthy and combined OA cohorts. This identified 18 proteins that were differentially expressed in healthy and OA patients, and individual sensitivity and specificity analysis was conducted on each after determining cut-off values. Although a multivariate supervised classification method would have been more appropriate and arguably simpler, the 18 proteins potentially represent a diagnostic fingerprint. It is interesting to note that none of the proteins in the biological panel of the current investigation were present. This highlights a potential limitation of some proteomic methods, i.e. polyacrylamide gel electrophoresis (PAGE) based mass spectroscopy, that may not be well suited to detection of small mediators such as cytokines. Furthermore, the excellent sensitivity of these techniques can itself pose a problem. For example, skin matrix components acquired during sampling often

need be excluded from the analysis. These studies highlight the potential utility of proteomics to biomarkers discovery. However, the specialist equipment, computational requirements and protein identification methods are arguably not currently accessible to most laboratories.

At the time of writing, this study is unique to the OA biomarker literature in applying the PLS-DA supervised multivariate statistical approach for pattern recognition and classification in wide-spectrum immunoassay analysis of biological fluids. PLS modelling assumes that the investigated system or process is actually influenced by just a few underlying latent variables. Both X- and Y-variables are assumed to be realisations of these underlying latent variables, and are hence not assumed to be independent. Therefore PLS is philosophically suited for the modelling of biological data. Partial least squares (PLS) and related methods inherently lend themselves to biomarker discovery where complex data sets have many variables, relatively few patients and the potential for missing values. Information is extracted by examining all data simultaneously and is robust to correlation between variables and noise in both X and Y variables [392]. PLS methods are frequently used in agriculture, environmental sciences, physical sciences and other biological fields. Medical disciplines commonly using the approach include oncology, neuroscience and microbiology especially when metabolomic and proteomic techniques are employed. A full documentation and discussion of these studies is beyond the scope of this thesis.

PLS and related MVDA methods are not completely alien to the field of OA. Their use has focussed on non-destructive evaluation of articular cartilage using near infrared and Fourier transform infrared imaging [491, 492, 493, 494]. Biomarker related studies using PLS methods are few and far between. Lamers et al. conducted a metabolomics study using ^{1}H nuclear magnetic resonance (NMR) spectroscopy and subsequent multivariate data analysis to define a urine metabolite fingerprint to discriminate and predict between patients with knee or hip OA and non-OA controls [495]. The same authors had previously piloted a similar study in OA-prone (Hartley) and OA-resistant guinea pigs [496]. The MVDA method used in both studies was PCA discriminant analysis, a technique related to PLS-DA. Furthermore, PLS regression demonstrated that this fingerprint correlated with radiographic severity in the human study. However,

the authors concede that identification of the NMR signals of the fingerprint would require extensive mass spectroscopy. A recent study by Gierman et al. used liquid chromatography tandem mass spectroscopy (LC/MS) on infra-patella (Hoffa's) fat pad explants followed by PLS-DA to identify a fingerprint of fatty acids and their oxygenated derivatives that could accurately differentiate end-stage knee OA and healthy donors [497]. Single molecule cohort comparisons had no diagnostic utility, and like the present study highlight the considerable potential of MVDA methods in biomarker research.

Implications

Foremost, this study is a proof-of-concept for the successful combination of wide-spectrum SF analysis and MVDA, an "immunomics" approach, for OA biomarker discovery. Potential avenues for future work that this study hopes to stimulate will be discussed further in Chapter 11.

This work highlights the need to reconsider and extend the concept of a biomarker to include multivariate approaches. Conceptually, the biological fingerprint for end-stage knee OA in this study is not a biomarker in the traditional sense. The measurement of markers within the defined fingerprints have limited utility on their own. There are no concentration reference ranges or cut-offs. The biomarker is the predictive multivariate model including existing patient (training) data, which together can be considered a calibrated instrument. In the case of the streamlined model, SF measurements of the 8 analytes (PIIANP, TIMP-1, ADAMTS-4, MCP-1, IP-10 and TGF-β3) from an unknown patient are evaluated by the existing model as test data. Based on the pattern i.e. internal correlation structure of the 8 markers, the model (biomarker) will diagnose end-stage OA, knee injury, inflammatory arthritis or "unknown". This is analogous to face or fingerprint recognition software when the structure or pattern of a person's face or fingerprint is matched to an existing database.

The present findings provide a highly accurate, reliable and specific combination biomarker and diagnostic tool for end-stage knee OA. The multivariate model (biomarker) is linked to a definite clinical endpoint since SF samples were taken at the time of

arthroplasty. Furthermore, the terminal joint failure of end-stage knee OA represents a convergence of clinical and structural endpoints. Therefore, the multivariate biomarker is potentially a surrogate endpoint for DMOAD development. Joint replacement surgery is often considered the ultimate clinical endpoint for knee OA and is in reality itself a surrogate for end-stage OA. However, it is an inaccurate endpoint that is influenced by numerous confounding factors unrelated to symptoms, structural pathology or treatment. A full discussion of these disparities is beyond the scope of this thesis, but include factors related to the patient (i.e. preferences, expectations, culture, socioeconomic status); the surgeon (i.e. preferences, experience, technical ability); and the healthcare framework (i.e. access and provision of resources, treatment guidelines) [95, 498, 203, 204]. Although all patients having joint replacement (should) have end-stage knee OA, not all patients with end-stage knee OA will therefore undergo joint replacement. Given the natural history of OA can be prolonged, arthroplasty can be a distant endpoint for patients with early OA who are potentially receptive to disease modification. Therefore, the utility of joint replacement surgery as an endpoint for disease-modifying treatment development is limited and trials would require large populations followed over long durations. There has been a largely unsuccessful initiative to define a "virtual joint replacement" (VJR) as an endpoint for DMOAD trials to eliminate many of the biases associated with arthroplasty as a study outcome while maintaining its easy dichotomous, OA-specific outcome [203, 499]. The multivariate end-stage knee OA biomarker models have the required attributes to qualify as a VJR. However, surrogate end-point status requires confirmation through rigorous biomarker qualification. Even if the this is not achieved, the present biomarker remains a useful adjunct to current methods of diagnosis and staging to guide patient management and assess current treatment modalities.

Limitations

There are a number of limitations to this study that need to be acknowledged. Like most human SF studies, the numbers of patients are relatively small especially in the reference groups. This is partly a function of the comparatively invasive nature of SF sampling. In particular, the collection of normal control synovial fluid was

unsuccessful and eventually abandoned despite adequate patient recruitment and ultrasound-guided methods. This has and will continue to be an obstacle for biomarker discovery since the truly normal knee and even early symptomatic OA do not have an effusion thus preventing satisfactory SF aspiration without lavage. The relatively small study population at a single location may limit wider geographical application of the multivariate biomarker and prevent surrogate endpoint and VJR status.

Both the knee injury and inflammatory arthritis cohorts represent heterogeneous groups of patients [Appendix F.3]. Larger more homogenous groups and/or subgrouping may have offered further information on these cohorts, but this was not the focus of the study or thesis. Knee injury patients were sampled at variable intervals after their index injury but there were no significant correlations between interval and marker concentrations [Appendix F.3]. There was limited clinical and structural information available for the inflammatory cohort of patients with RA or psoriatic arthritis who were on a range of treatments. In particular, the existence of OA secondary to inflammatory arthritis was not excluded. However, the multivariate approach was robust to the heterogeneity of the comparator groups.

This study had very stringent inclusion and exclusion criteria for OA cohort membership, but the potential confounding effects of medications were not considered. In particular, knee injury patients may not have been on (m)any medications, whereas end-stage knee OA patients were likely to be taking a combination of paracetamol, NSAIDs and opioids, and patients with inflammatory arthritis were taking corticosteroids and/or disease-modifying drugs potentially in addition to analgesia. Patient age, which differed significantly between cohorts, may also represent a confounding factor in the multivariate models. Ideally one would study age-matched patient groups, but this would have been difficult to achieve since advancing age is inextricably linked to OA development and knee injuries tend to occur in young adults.

9.6 Conclusions

Null Hypothesis: Wide-spectrum SF analysis cannot provide a combination biomarker ("biological fingerprint") for end-stage knee OA. ***- Rejected***

This chapter has been successful in addressing the ultimate aim of the thesis to discover a biological fingerprint for end-stage OA. Wide-spectrum immunoassay analysis using multiplex platforms was combined with a PLS-DA multivariate statistical analysis. The novelty of the approach was in linking previously described analytical and statistical methods with the goal of biomarker discovery. This strategy can be described as "multiplex immunomics". The quantitative multivariate models incorporating biological fingerprints demonstrated excellent diagnostic accuracy, reliability, specificity and external validity for end-stage knee OA. The findings and approach represent a new breed of biomarker and potential diagnostic tool to complement existing end-stage knee OA diagnostic criteria, stage disease, define inclusion/exclusion criteria for therapy trails and monitor the efficacy of such interventions. This work represents a proof-of-concept that will initiate and stimulate further investigations that can drive OA biomarker research into a new era to facilitate the development of disease-modifying OA therapies.

The findings of this study have resulted in a patent application:

Isis Innovation Ltd. *Methods and System for Determining the OA Disease Status of a Subject*. GB Application No: 1404634.6. (Filed 14th March 2014)

Chapter 10

Thesis Discussion

"There is a fine line between fishing and just standing on the shore like an idiot"

- Steven Wright (American comedian)

Introduction

The lack of disease-modifying pharmacological therapies for osteoarthritis (OA) has been linked to a lack of suitable biomarkers. When starting work on this thesis there were no reliable, quantifiable and easily measured biomarkers for OA that enable diagnosis, prognosis or monitoring of treatment effect at the individual level. It is now generally accepted that a condition as complex as OA will be best served by combinations of multiple biomarkers. There are a variety of biological fluids that are potential sources of biochemical biomarkers for OA. Attempts at biomarker discovery have largely concentrated on biological fluids distant from the synovial joint such as peripheral blood and urine, which have many limitations. However, synovial fluid (SF) best reflects the biological milieu of the joint and offers a direct measure of joint pathophysiology.

This thesis set out to discover a combination biomarker or "fingerprint" for knee OA through wide-spectrum SF analysis of physiologically relevant cytokines, chemokines, growth factors, matrix enzymes and markers of cartilage turnover. A series of studies were performed to first ensure the reliability of multiplex immunoassay technologies

and then utilise these findings for biomarker discovery in patients with end-stage knee OA. Detailed discussions specific to each study are given in their respective chapters. Here, an overview of the thesis is provided.

Some may view the work presented in this thesis as a "fishing expedition" and in many respects this may be an accurate sentiment. However, a successful fishing expedition requires the correct bait, hooks, rods and nets and knowledge of where best to fish to maximise the chances of a good catch or haul. To this end, this work represents a fruitful fishing expedition.

Part I - Methods Development and Practical Implications

Osteoarthritic joints often do not offer sufficient volumes to allow analysis of multiple markers in a single sample by conventional methods such as ELISA. Furthermore, SF is challenging to investigate due to its complex matrix and viscous non-Newtonian nature. Despite these obstacles, there is a growing minority of studies focusing on SF analysis partly stimulated by the growth of multiplex immunoassay and "-omic" technologies. In particular, the Luminex xMAP and Meso Scale Discovery (MSD) enable multiplex immunoassay analysis in small ($\leq 50\mu l$) sample volumes, whilst also increasing cost and time efficiency. The commercial protocols for these platforms are validated only for serum, plasma and culture supernatant. The first part of this thesis aimed to optimise and validate SF analysis by the Luminex and MSD multiplex platforms (Chapters 3 and 4), and address potential practical requirements of SF sampling (Chapters 5 and 6). Chapter 2 outlined the general methods, materials and equipment used for the forthcoming studies. Although not within the specific remit of this thesis, the work in Chapters 3 and 4 highlighted the strict procedures for immunoassay data processing and quality control that are essential yet frequently neglected in biomarker studies.

Chapter 3: Synovial Fluid Treatment

Assay precision is important for biomarker discovery, particularly when characterising and comparing patient cohorts, and also an essential part of assay validation. This

study aimed to ascertain whether SF treatment by dilution or hyaluronidase (HAse) could improve intra-assay precision on the Luminex and MSD multiplex platforms. The data demonstrated that SF HAse treatment permits better precision during multiplex analysis on the Luminex platform compared to neat or diluted samples. Furthermore, measured concentrations were higher and recovery was elevated after HAse treatment. These performance advantages may be a result of increased efficiency of bead-based sandwich immunoassays in the fluid phase. One can, therefore, recommend routine HAse treatment of SF samples prior to analysis on the Luminex platform. Hyaluronidase treatment conferred no conclusive benefits to assay precision or signal on the solid-phase MSD platform, which thus seems to be more resistant to matrix effects. Wide-spectrum characterisation of low volume SF samples from OA patients will likely require analysis on multiple platforms to gain the maximum biological information. Therefore, it is possible to advocate routine HAse treatment to reduce SF sample viscosity for biomarker research in OA.

Chapter 4: Multiplex Immunoassay Platforms

Neither of the two dominant multiplex immunoassay technologies, Luminex and MSD, has been validated for SF analysis. The cross-comparison of multiplex formats and comparison against the existing gold standard of ELISA are required to facilitate quality control and guide optimum assay selection for biomarker discovery. Previous comparison studies in culture medium and serum/plasma cannot be extrapolated to the complex biological matrix of SF. This chapter made a detailed comparison between Luminex and MSD platforms during the analysis of real clinical SF samples from end-stage knee OA. This comprehensive study goes some distance in validating the Luminex and MSD platforms for SF analysis with both platforms demonstrating acceptable precision and accuracy. The MSD platform showed lower analytical noise (LOD) and better quantification limits than the Luminex platform. It was, therefore, better able to detect and quantify low-level analytes (i.e. IL-1β & TNF-α) in OA. The MSD platform has a similar protocol to the familiar ELISA and a rapid analyser read time. At the time of performing the experimental work in this thesis MSD multiplex assays were limited to a maximum of 10 markers, but recent platform evolution now allows up

to 100-panel assays. The Luminex system allows high multiplexing, but data is very dependent on bead events and the machine read time can thus often be long. The Luminex analyser requires regular careful maintenance and calibration. Through better bead retention, the magnetic Luminex system minimises potential wastage of precious SF samples and is to be recommended over its polystyrene counterpart.

Cross-platform comparisons of cytokine measurements suggest that biomarker measurements are at best only quasi-quantitative, with poor absolute agreement but excellent relative agreement. Therefore comparison of marker concentrations between studies needs to be made with caution. The reason for poor absolute agreement is most likely due to differences in antibody pairs and kinetics specific to the technology and precise assay. Once an analytical method has been used to measure a marker in a study, there needs to be strict consistency in methodology to facilitate evaluation and comparison of data. This should ideally encompass sample collection, storage and preparation; assay platform, manufacturer, specific kit, reagents and calibrator lot numbers. Biomarker studies need be consistent and explicit regarding assays.

Chapter 5: Anatomical Site of Sampling

Having established the analytical requirements for reliable SF analysis in Chapters 3 and 4, the importance of consistency in sample collection was raised. The flow of SF during knee motion is not simple and despite the free communication of SF between compartments of the knee joint, there is evidence in the literature for "physiological compartmentation" of the synovial space. In addition, SF has complex rheological properties and both its flow and diffusion are difficult to model. These factors raise the possibility of compartment-specific SF biology. Anatomical variations in SF analysis would have important practical implications for the consistency of sampling needed for biomarker research, especially for unicompartmental OA i.e. anteromedial knee OA (AMOA) where there is macroscopically preserved lateral compartment cartilage with limited osteophytosis or synovitis. The objective of this study was to determine if the anatomical site of harvest affects SF measurements in patients with end-stage AMOA and tri-compartmental OA (TCOA). SF samples were drawn simultaneously from widely separated medial, lateral and patello-femoral compartments. The data

demonstrated excellent agreement between compartment-specific SF measurements of 22 cytokines, chemokines and TGF-β isoforms, irrespective of patient groups. Therefore, the anatomical site of sampling does not affect SF analysis even when there are significant compartmental differences in structural burden of disease. Although the number of markers assessed is by no means exhaustive, the findings strongly suggest biological equilibrium in SF composition between joint compartments in end-stage knee OA and consequently the site of SF collection is not an important practical implication.

Chapter 6: Short-term Temporal Stability

Short-term fluctuations in biomarker measurements unrelated to disease progression or treatment would invalidate the utility of biomarkers for patient assessment. The SF biology of end-stage knee OA should reflect the biological failure associated with the extent of clinical and structural (radiographic) severity. Consequently, one would expect SF measurements to be stable over short periods of time during which there is no further clinical or structural deterioration. This study aimed to formally test this assumption. The results showed strong to excellent agreement in SF measurements of 22 cytokines, chemokines and TGF-β isoforms taken about 2 weeks apart in patients with end-stage knee OA. This suggests biological stability in the SF over short intervals during which there has been no change in disease severity. There was some evidence of analyte-specific intra-individual variability, but no overall directional change between time-points. However, one must again be cautious in generalising the results to all possible SF markers. Ideally, potential biomarkers highlighted by exploratory studies should be assessed for short-term sampling stability as part of advanced validation. Furthermore, any variability should be used to define minimally important biological and clinical changes.

Part II - Wide-spectrum Synovial Fluid Analysis

The second part of this thesis utilised the findings of Part I to explore the SF biology of end-stage knee OA with a view to biomarker discovery. End-stage knee OA was

chosen not only because of the accessibility of clinical samples during arthroplasty. By definition, end-stage knee OA represents the only known true convergence of clinical symptoms and structural pathology in OA. A detailed description of the study design and methodology used to answer the research questions was provided in Chapter 7. It is worth highlighting the detailed and stringent criteria for cohort membership and the difficulty in obtaining SF from asymptomatic "normal" knees resulting in recruitment for this cohort being abandoned. The challenges of SF measurements below assay quantification and/or detection limits (left censored data) was introduced in Part I and again many studies neglect or inappropriately handle such data resulting in significant bias. A simple and robust method (β-substitution), previously shown to be as effective as more complex gold-standard statistical techniques, was chosen for the thesis and at the time of writing was the first reported use in the OA biomarker literature.

Chapter 8: Structural Phenotypes of End-stage Knee OA

Knee OA is a heterogenous syndrome with evidence of distinct structural phenotypes on the basis of regional articular cartilage degeneration, subchondral bone remodelling and synovitis. In particular, AMOA is the most common structural phenotype of knee OA with a highly consistent topographical pattern of disease that is reflected in the biology of articular cartilage. The work in this chapter attempted to discover burden of disease biomarkers by assessing for potential differences in the SF biology of 60 well characterised patients with cumulatively more severe structural phenotypes of end-stage knee OA: AMOA, AMOA plus patello-femoral OA and TCOA. All patients were well matched for age, BMI and clinical severity as measured by the Oxford Knee Score. This novel approach, using both standard and multivariate statistical methods, surprisingly revealed no phenotypic differences in the concentration or variability of a wide range of 35 cytokines, chemokines, growth factors, matrix enzymes and markers of matrix turnover despite significant differences in structural burden of disease. This raised a challenging problem: Why does focal OA confined to one joint compartment have an equivalent SF biology to extensive 3-compartment degeneration for a wide range of markers implicated in OA pathophysiology? The work in Part I of the thesis and internal validation of all measurements using samples from patients with knee injury

or inflammatory arthritis rules out analytical artefact. Potential coexisting explanations for the chapter findings that each have support from the literature include:

- a “ceiling phenomenon” where SF biological activity reaches saturation at earlier stages of disease due to pre-macroscopic OA changes in preserved joint compartments.
- a “burnout phenomenon” where SF biological activity falls as progressively more articular volume is lost with cumulative compartment involvement.
- a “clearance phenomenon” where there is greater SF clearance of markers due to greater (unaccounted) synovitis associated with more extensive structural disease.

The latter highlights the principles of SF kinetics and clearance that are well established, but are under-appreciated despite significant implications for the interpretation of concentration measurements. Nonetheless, the results suggest that end-stage knee OA, defined as clinically severe and radiologically advanced disease that is refractory to non-operative management, has a constant SF biology irrespective of the “volumetric” extent of cartilage degeneration. The SF biology of end-stage knee OA can therefore provide a useful terminal benchmark for OA biomarker research.

Chapter 9: Synovial Fluid Fingerprinting of End-stage Knee OA

This final experimental chapter represents the culmination of the thesis by returning to its ultimate aim, to discover a combination biomarker or “biological fingerprint” for OA, using end-stage knee OA as a stable (Chapter 6) and terminal (Chapter 8) clinical, structural and biological entity. Wide-spectrum SF immunoassay analysis of patients with end-stage knee OA, knee injury and inflammatory arthritis using a mixture of multiplex and ELISA platforms was combined with PLS-DA multivariate statistical analysis.

Multivariate models with biological fingerprints for end-stage OA were derived from 35 pathophysiologically relevant cytokines, chemokines, growth factors, matrix enzymes

and markers of cartilage turnover measured in SF. An initial full model from 20 analytes that qualified for quantitative analysis was simplified to an 8-marker streamlined model with an end-stage knee OA fingerprint comprising "ridges" of PIIANP, TIMP-1, ADAMTS-4 and MCP-1 with "troughs" of IP-10 and TGF-β3. Both models achieved excellent (>95%) diagnostic accuracy, reliability and specificity, and were successfully validated with 10 new end-stage knee OA test samples. Although the multivariate model from 12 analytes that did not qualify for quantitative analysis were suboptimal, they revealed markers such as MMP-13 and IL-1RA that loaded favourably to end-stage knee OA. These markers should be reassessed with more sensitive assays and could potentially refine the existing biological fingerprints. The multivariate models demonstrated relative proximity of knee injury patients, without any macroscopic evidence of OA, to those with end-stage knee OA. This is consistent with growing evidence that knee injury presents a high risk for OA development.

This study represents a potential landmark in OA biomarker discovery for both the end result of a diagnostic biomarker fingerprint for end-stage knee OA and for the multivariate approach used in deriving it. The composition of the biological fingerprints were rationalised and intellectually validated with references to the existing OA literature. The streamlined model and fingerprint have potential for industry development as a biological tool i.e. an OA-plex to complement existing OA diagnostic criteria, stage disease, define inclusion/exclusion criteria for therapy trails and monitor the efficacy of such interventions. The benefits of the multivariate approach were immediately apparent, in that multiple between-cohort comparisons offered limited useful conclusions barring the finding that elevated PIIANP and reduced TGF-β isoforms may be specific to end-stage knee OA. The inherent properties of the PLS regression make it ideally suited to biological systems and biomarker research. Pattern recognition is achieved through simultaneous analysis of correlation structure between analyte measurements. This is particularly relevant given the quasi-quantitative nature of immunoassay measurements described in Chapter 4. Furthermore, the inclusion of 2 non-OA comparator groups, one on either side of the OA disease spectrum, significantly increased the specificity of potential biomarkers and boosted the statistical power from analysis of relatively few patients.

The work in this chapter represents a proof-of-concept that hopes to upgrade current concepts of biomarkers and stimulate a new wave of OA biomarker research.

Chapter 11

Directions for Future Work

One cannot go through the process of writing a thesis without highlighting areas for future work. Further investigations have been recommended during the course of individual study and chapter discussions. However, this chapter focusses on proposals for future research stimulated by the successful discovery of a diagnostic biomarker fingerprint for end-stage knee OA.

It is worth stating again that the biomarker is the predictive multivariate model including existing patient data. The most immediate and essential priority is further validation. Much of the analytical validation has been conducted in Part I of the thesis and partial external validation has been achieved using 10 new patients with end-stage lateral compartmental knee OA. However, full external validation of the multivariate models and biological fingerprints requires SF analysis of new patients with knee injury and inflammatory arthritis. It would also be reassuring to test more end-stage knee OA patients including the other structural phenotypes described in Chapter 8. The data generated from validation patients can be subsequently incorporated into the existing models to increase robustness and may offer refinement of the biological fingerprints.

The OA biomarker fingerprint was generated from SF samples harvested from patients at the time of arthroplasty, and therefore the converged clinical and structural state of end-stage knee OA. This makes it a potential candidate for a surrogate endpoint for end-stage knee OA as a “virtual knee replacement”. However, this requires a rigorous process of biomarker qualification. Ideally both validation and qualification should be conducted using identical assays that generated the multivariate biomarker models.

This will be challenging in terms of the number of different assays, platforms and manufacturers used. Even the current streamlined 8-biomarker model requires one MSD assay, 3 ELISAs and 3 Luminex assays. Furthermore, the use of multiple assays and in particular ELISA increases the required SF volumes. Therefore, the immediate validation work should occur in parallel with the development of an "OA-plex" on a single platform to maximise efficiency and minimise time to translation. One of the great advantages of the PLS multivariate methods is that pattern recognition is based on the internal correlation structure of analyte measurements and not the absolute concentrations that are assay-specific. This presents the freedom and opportunity to develop a singe multiplex assay format through industry partnership. The MSD platform is an obvious candidate due its analytical sensitivity, robustness to matrix effects and the relative ease of custom immunoassay development. It is also worth considering alternatives such as the aptamer-based SOMAscan platform that is a proteomic technology that combines the advantages of immunoassays and microarrays. Adjacent research using SOMAscan would allow the measurement of more than 1000 analytes in the SF samples to discover further or better OA biomarker fingerprints, and in particular assess the importance of MMP-13 and IL-RA that showed promise in the categorical multivariate model. The development of a suitable single platform "OA-plex" would hugely facilitate biomarker qualification and any subsequent studies.

As a diagnostic tool, one can argue that the current biomarker model is redundant in the presence of clinical and radiographic measures indicative of end-stage knee OA. However, it provides a biological diagnosis that has been thus far unavailable. It would be of great interest to assess how this biological diagnosis of end-stage knee OA is affected by interventions such as intra-articular corticosteriod injection or joint replacement. In particular, a persistent biological diagnosis of end-stage knee OA following arthroplasty, with alleviated pain and functional improvement, would be difficult to reconcile and may derail biomarker qualification.

It has been assumed that the present biomarker model is specific for only end-stage knee OA, but this cannot be confirmed without assessing earlier stages of disease. It is not known exactly when the current OA fingerprint first appears along the continuum from normal joint to joint death. It has already been seen that end-stage knee OA has a

constant SF biology irrespective of the "volumetric" extent of cartilage degeneration (Chapter 8). It is therefore not inconceivable that the present OA fingerprint and the multivariate biomarker model may diagnose OA in patients prior to terminal joint failure and arthroplasty. This would preclude surrogate endpoint status as a "virtual knee replacement", but may be potentially more useful for development of disease-modifying therapies. Serial longitudinal analysis of large well-characterised cohorts with associated clinical, radiological and other biological measurements would be the gold standard approach to answering this question. Such cohorts exist in the form of the Osteoarthritis Initiative (OAI) and Cohort hip and Cohort knee (CHECK), but SF sampling is not currently a feature of either of these commendable research enterprises. The invasive nature of sample collection and potential low volume yields is an obvious challenge. It is imperative that SF collection in these large longitudinal cohorts should be actively encouraged, even if recruitment and successful harvest is only possible in a small fraction of patients. In the absence of prospective longitudinal samples, an alternative short-term solution is the cross-sectional analysis of well-characterised patient cohorts at sequentially earlier stages of clinical and structural disease. In addition, increasing efforts and innovations must be made to obtain SF from patients with normal asymptomatic joints. Such work would reveal biomarker fingerprints for the earlier OA disease stages and normal joints, which would have far reaching implications for the development of disease-modifying therapies. The multivariate approach allows patient cohorts and their biological information across the OA continuum, from normal to end-stage knee OA, to be combined to refine their respective fingerprints and generate an extended model that can both diagnose and monitor progression. However, it remains important that SF samples from all patients used to generate such a multivariate biomarker model should be analysed for the same SF markers in the same way, whether by multiple different assays as in this thesis or a newly developed single "OA-plex".

This thesis was motivated by a general feeling of pessimism due to a lack of OA biomarkers, largely attributed to narrow-spectrum analysis of peripheral biological fluids. It therefore hopes to "reboot" OA biomarker discovery. Wide-spectrum SF analysis produced a highly accurate and reliable diagnostic multivariate biomarker

model with a biological fingerprint for end-stage knee OA. The results and multivariate approach serve as a proof-of-concept to stimulate a new era in OA biomarker discovery.

Appendix

Appendix A

Immunoassay Principles and Performance

A.1 Immunoassay Platforms

Immunoassays are bioanalytical methods that measure the concentration of target molecules (analytes) in solution through the use of specific immunoglobulin antibodies. A quantifiable signal is produced to indicate the binding of the antibody to its antigenic epitope on the analyte. This is achieved by linking the formation of immune complexes to a detectable labelling system, for example an enzyme reaction, which produces a detectable signal i.e colorimetric change, fluorescence or luminescence. Analysis requires construction of a standard curve, which represents measured signal intensity as a function of the analyte concentration, using a series of known analyte concentrations (calibrators). Unknown analyte concentrations are then determined from this standard curve.

Enzyme-linked Immunosorbent Assay (ELISA)

As the name suggests, ELISA methods use an enzyme reaction for detection of the immunoassay that occurs on a solid surface. ELISAs are typically performed in 96-well (or 384-well) polystyrene plates, which will passively bind antibodies and proteins.

Analyte molecules within a sample are immobilised onto a solid surface, usually the floor of a polystyrene plate. Analyte capture can be non-specific by adsorption onto the

surface, or specific by an analyte-specific capture antibody previously adsorbed onto the surface (i.e. “sandwich” ELISA). Captured analyte molecules are then detected by an analyte-specific antibody against a different epitope than the capture antibody. The detection antibody can be conjugated to an enzyme (i.e. direct ELISA), or can itself be detected by a secondary antibody that is conjugated to an enzyme (i.e. indirect ELISA). The assay plate is typically washed with a mild detergent solution between steps to remove excess sample, or non-specifically bound proteins and antibodies. The final step involves addition of the enzyme’s substrate to produce a detectable signal, that has traditionally been a measurable colour change in the substrate.

Luminex xMAP

Luminex xMAP is a fluorescent bead-based immunoassay technology that allows the simultaneous measurement of up to 100 different analytes within a single sample. Conceptually, the method can be thought of as a hybrid of sandwich ELISA and flow cytometry.

The technique utilises microspheres (“beads”) encoded with internal red and infrared fluorescent dyes at distinct ratios. Each unique population of microspheres is given an identification number called a bead region, and can be conjugated to different analyte-specific capture antibodies. These antibody bead classes or sets provide a large surface area for analyte capture. Captured analyte molecules are then detected by analyte-specific detection antibodies that are linked to biotin. The sandwich immunoassay is then reported by streptavidin conjugated to the fluorescent protein, R-phycoerythrin (RPE) that binds non-specifically to biotinylated detection antibodies. A mixture of different antibody bead classes and corresponding detection antibodies enables the simultaneous multiplex detection of many analytes in the same sample.

Luminex assays are performed in 96-well microtitre plates. During data acquisition the contents of each plate well are drawn into a flow cytometry/laser excitation–based analyser. Precision fluidics align the beads in single file through a flow cell where two lasers simultaneously excite each bead individually. A red (635nm solid state) classification laser excites the internal dyes in each bead, identifying its bead region

and therefore the target analyte. A green (532nm Nd:YAG) reporter laser excites the RPE reporter molecule associated with the bead's sandwich immunoassay complex, which allows quantitation of the captured analyte.

There are two types of Luminex microspheres: the polystyrene only (5.6μm) bead and the magnetite-coated (6.5μm) bead. Nonmagnetic beads utilise a vacuum workflow requiring filter plates and vacuum filtration for wash steps. Magnetic beads can use a magnetic workflow as well as vacuum workflows. The magnetic workflow allows automated dispense and aspiration washing with robotic wash stations that use magnets to retain beads.

MesoScale Discovery

MesoScale Discovery (MSD) technology uses an electrochemiluminescence (ECL) detection system that allows the simultaneous measurement of up to 100 different analytes within a single sample. The basic principles and workflow, however, mirror that for a direct sandwich ELISA.

The MSD platform typically uses specialised 96-well multi-array plates with up to 10 flat carbon electrodes ("spots") per well. There are also 384-well plates with up to 4 spots per well, and 24-well plates with up to 100 spots per well. Each electrode or "spot" is coated with a different analyte-specific capture antibody. In contrast to ELISA, captured analytes on the spot electrodes are detected by an analyte-specific SULFO-TAG labelled detection antibody. The SULFO-TAG label is an electrochemiluminescent compound containing ruthenium. Addition of a read buffer provides an optimum chemical environment for the ECL reaction.

The detection process is initiated by voltage applied to the spot electrodes at the base of the plate. This causes the SULFO-TAG label to emit light at a wavelength of 620nm. The read buffer contain coreactants that enhance the ECL signals, and are stimulated when in proximity to the spot electrodes. Multiple excitation cycles of each label amplify the signal to enhance light levels and improve sensitivity. Background signals are minimal because the stimulation mechanism (electricity) is decoupled from the signal (light).

Only labels near the spot electrode are excited and detected allowing the concentration of the analyte to be determined relative to the particular spot electrode.

A.2 Immunoassay Performance Parameters

Bead Events

In the Luminex system, a *bead region* (or *bead set*) describes beads with the same spectral signature of internal red and infrared dyes, coated with a specific capture antibody. A *bead event* is defined as a bead bound to its specific ligand and corresponding detection antibody. The analyser counts the number of these complete sandwich immunoassays. Thirty-five bead events (for each unique bead set) is reported as the minimum requirement for statistically significant analyses [259]. The recommended minimum bead events/region varies depending on the assay manufacturer. For example Bio-Rad and R&D Systems suggest 50 bead events/region; Life Technologies suggest 100 bead events/region.

Limit of Detection

The limit of detection (LOD) is the lowest concentration that can be reliably distinguished from blank measurements and is therefore qualitatively present in the sample [500, 501]. The formula used to calculate the background-subtracted signal with 95% confidence of being greater than the blank signal (LOD_{net}) is given by:

$$LOD_{net} = 4.65\sigma_b$$

where σ_b is the standard deviation (SD) of blank measurements [502]. The LOD concentration is then calculated from log-linear regression of the 4 lowest calibrator point concentrations plotted against their net signal. In the event that $\sigma_b = 0$, the LOD concentration is derived from the uncorrected signal with 95% confidence of being greater than the blank (LOD_{raw}):

$$LOD_{raw} = \bar{x}_b + 3.29\sigma_b$$

$$LOD_{raw} = \bar{x}_b \; (if \; \sigma_b = 0)$$

where $\bar{x}_b$ is the mean blank signal [502]. Log-linear regression was conducted as above using uncorrected signals. These LOD calculations are more stringent than usually used in the literature i.e. blank signal plus 2, 2.5 or 3 SDs.

Individual LODs for each analyte were recorded from every assay plate that was run during the course of the thesis. The average LOD for a particular assay was described by the geometric mean from the multiple plate runs. These average LODs [Appendix C] were then applied to the analyte concentrations from patient samples.

Limits of Quantification

The (lower) limit of quantification (LOQ) is the lowest concentration of analyte that can be measured with appropriate accuracy and reliability [501, 500]. By definition the LOQ must be greater than or equal to the LOD. For this thesis, the LOQ was defined from the standard curve as the lowest calibrator point, greater than or equal to the LOD, with both a back-calculated concentration $\pm$20% of nominal and replicate coefficient of variation $\leq$20% [260, 261, 501, 500].

The upper limit of quantification (ULOQ) is the highest concentration of analyte that can be measured with appropriate accuracy and reliability [501, 500]. The ULOQ was defined from the standard curve as the highest calibrator point with both back-calculated concentrations $\pm$20% of nominal and replicate coefficient of variation $\leq$20% [260, 261, 501, 500].

Individual LOQs and ULOQs for each analyte were recorded from every assay plate that was run during the course of the thesis. The average LOQ and ULOQ for a particular assay was described by the median from the multiple plate runs. The median was used because the calibrator points of the standard curve are arbitrary discrete concentrations. These average limits [Appendix C] were then applied to the analyte concentrations from patient samples.

Intra-assay Precision

Precision is the degree of agreement among individual test results when the analytical method is repeated for a sample. The intra-assay precision was determined by the coefficient of variation (CV) for measured replicate concentrations, given by the formula:

$$CV = \frac{\sigma}{\bar{x}}$$

where $\bar{x}$ is the mean replicate concentration and σ is the standard deviation. This was expressed as a percentage. A CV $\leq$20% is the generally accepted criterion for satisfactory intra-assay precision, unless replicate concentrations fall on the lower tail of the standard curve when 25% is the threshold [260, 261, 262, 263].

Recovery

Accuracy is the closeness of a test result obtained by the analytical method to the true (known) value. *Recovery* is a measure of assay accuracy, and is defined as the fraction of the analyte determined in a known amount of the analyte. It is usually expressed as a percentage and the general formula is

$$Recovery\ (\%) = \frac{C_{obs}}{C_{exp}} \times 100$$

where C_{obs} is the observed or measured concentration and C_{exp} is the expected or known concentration. The *relative error* (RE) is a related parameter of accuracy, also expressed as a percentage, given by:

$$RE\ (\%) = \frac{C_{obs} - C_{exp}}{C_{exp}} \times 100 \qquad \therefore\ RE = Recovery - 1$$

Recoveries between 80% and 120% (i.e. RE within ±20% of expected) were considered acceptable [260, 261, 262].

Matrix Interference

The physical properties and constituents of a sample can potentially influence the measurement of analytes within it. These are know as matrix effects. *Spike recovery* and *dilution recovery* assess sample matrix interference on assay accuracy. In particular, its is used to determine if the assay is affected by differences between the diluent used to prepare the standard curve and the sample matrix.

Spike Recovery

For spike recovery, a sample is spiked with known concentrations of analyte i.e. calibrator from the standard curve, and the recovery is given by the formula:

$$Spike\ Recovery\ (\%) = \frac{C_{obs} - C_{nat}}{C_{spike}} \times 100$$

where C_{obs} is the final measured assay concentration in the spiked sample, C_{nat} is the measured concentration for the native unspiked sample and C_{spike} is the expected known spike concentration.

Dilution Recovery

Dilution recovery is determined from a series of increasingly diluted samples, and is given by the formula:

$$Dilution\ Recovery\ (\%) = \frac{C_{obs} \times D_f}{C_{exp}} \times 100$$

where C_{obs} is the final measured assay concentration, C_{exp} is the expected (neat) concentration, and D_f is the dilution factor.

Both spike and dilution recovery are also measures of assay parallelism, to assess the validity of the calibrator diluent as a substitute sample matrix for standard curve generation [503].

Appendix B

Materials and Equipment

B.1 Part I Immunoassays

Immunoassays used in Part I of the thesis are given in Table B.1. All immunoassays were conducted according to the manufacturer's protocol by a single researcher (Chethan Jayadev).

B.2 Part II Immunoassays

Immunoassays used in Part II of the thesis are given in Table B.2.

Oxford - GlaxoSmithKline Collaboration

Assays for matrix enzymes MMP-1, MMP-3, MMP-9, TIMP-1 and ADAMTS-4, and markers of cartilage turnover ARGS neo-epitope, COMP and PIIANP were conducted at Biopharm R&D, GlaxoSmithKine (GSK), Stevenage, UK. Synovial fluid samples were transferred to GSK, Stevenage on dry ice and immediately stored at -80°C. No sample underwent more than 2 freeze-thaw cycles before analysis at GSK. For any particular marker assayed, all samples analysed had undergone the same number of freeze-thaw cycles. All remaining assays were performed at the Botnar Research Centre, Oxford by a single researcher (Chethan Jayadev).

Assay	Manufacturer	Cat No.
ELISA		
Human IL-6 DuoSet	R&D Systems	DY206
Human TNF-α DuoSet	R&D Systems	DY201
Luminex		
Human Cytokine 10-plex ^{p} IL1-β, IL-6, TNF-α, IL-4, IL-8, IL-10, GM-CSF, IL-5, IFN-γ & IL-2	Life Technologies	LHC0001
Human Cytokine 25-plex ^{m} IL1-β, IL-1RA, IL-2, IL-2R, IL-4, IL-5, IL-6, IL-7 IL-8, IL-10, IL-12(p40), IL-13. IL-15, IL-17 TNF-α, IFN-α, IFN-γ, GM-CSF, MIP-1α, MIP-1β IP-10, MIG, Eotaxin, RANTES & MCP-1	Life Technologies	LHC0009M
Bio-Plex Pro Human Cytokine IL-6 ^{m} & Bio-Plex Pro Human Cytokine Standards Group I & Bio-Plex Pro Reagent Kit	Bio-Rad	171-B5006M 171-D50001 171-304070
Bio-Plex Pro Human TGF-β 3-plex ^{m} TGF-β1, TGF-β2 & TGF-β3	Bio-Rad	171-W4001M
MesoScale Discovery		
Human Proinflammatory-4 II Ultra-Sensitive IL1-β, IL-6, IL-8 & TNF-α	MesoScale Discovery	K15025C

Table B.1 – Immunoassays used in Part I of the thesis. Assays kit names, target analytes, manufacturers and catalogue numbers are shown. For Luminex assays, the *p*=polystyrene bead system, and *m*=magnetic bead system.

		Assay	Manufacturer	Cat No.
Luminex	*o*	Human Cytokine 25-plex ^{m} IL-2, IL-12, IL-15, GM-CSF, IL-1Ra, IL-4, IL-10, IL-2R, RANTES, MIP-1α, MIP-1β, MCP-1, IP-10, Eotaxin & MIG	Life Technologies	LHC0009M
	o	Bio-Plex Pro Human TGF-β 3-plex ^{m} TGF-β1, TGF-β2 & TGF-β3	Bio-Rad	171-W4001M
	o	VersaMap Human MMP-13 ^{p}	R&D systems	
MesoScale Discovery	*o*	Human Proinflammatory-4 II IL-1β, TNF-α, IL-6 & IL-8	MSD	K15025C
	g	Human MMP 3-plex MMP-1, MMP-3 & MMP-9	MSD	K15034A
	g	Human TIMP-1	MSD	K151JFC
	o	Prototype 4-plex BMP-2, BMP-7, LIGHT & DcR3	MSD	N45ZA-1
	g	^{374}ARGS Neo-epitope	GSK Custom in-house MSD	
ELISA	*g*	Human ADAMTS-4	CusaBio	CSB-EL0001311HU
	g	Human COMP	BioVendor	RD194080200
	g	Human PIIANP	Millipore	EZPIIANP-53K

Table B.2 – Immunoassays used in Part II of the thesis. Assays kit names, target analytes, manufacturers and catalogue numbers are shown. For Luminex assays, the *p*=polystyrene bead system, and *m*=magnetic bead system. SF analysis was conducted in Oxford (*o*) and by GSK (*g*).

B.3 Materials and Equipment

Tween-20, 1% bovine serum albumin (BSA) and bovine testicular hyaluronidase (type I-S, 618.4 U/mg) were purchased from Sigma. Sterile filtered PBS (10X) was purchased from Lonza and diluted as required.

96-well plates, reagents and wash solutions provided by the manufacturers were used for all immunoassays.

A BioTek ELx50 microplate washer (Oxford) or Molecular Devices SkanWasher 300 (GSK) was used to perform automated plate washing. An IKA MTS 2/4 digital microtitre plate shaker was used for incubation steps. For Luminex assays, a vacuum filtration manifold was used for the polystyrene bead assays using (1.2μm pore) filter-bottom plates, and the magnetic separation manifold was used for magnetic bead assays using flat-bottom plates. Assay plates were analysed by a Luminex xMAP-200 system, MSD Sector Imager 6000 or a Molecular Devices SpectraMax plate reader for ELISA plates.

ELISAs in Part I were conducted in Maxisorp 96-well plates purchased from Nunc. Immunoassay detection used the peroxidase substrate 2,2´-Azino-bis-(3-ethylbenzo-thiazoline-6-sulfonic acid) purchased from Sigma. The ELISA plates were analysed using a Molecular Devices SpectraMax Plus-384 plate reader. The absorbance was read at 450nm.

Appendix C

Limits of Detection and Quantification

The limit of detection (LOD) is described by the geometric mean and 95% confidence limits (95%CI). The lower (LOQ) and upper (ULOQ) limits of quantification is described by the median and interquartile range (IQR). The total number of assay plate runs conducted during the thesis work is provided.

Assay	LOD	95%CI	LOQ	IQR	ULOQ	IQR
IL-1β	3.33	1.66 - 6.68	7.6	3.80 - 11.39	8300	4150 - 8,300
IL-6	0.80	0.47 - 1.37	1.9	1.90 - 1.90	2766.7	1383.3 - 4150
IL-8	1.10	0.50 - 2.43	5.85	5.85 - 5.85	12800	6400 - 12800
TNF-α	1.62	1.29 - 2.03	3.98	3.98 - 3.98	8700	4350 - 8700
IL-10	1.62	1.22 - 2.16	10.01	10.01 - 10.01	7300	7300 - 21900
IL-13	6.04	3.20 - 11.40	9.28	9.28 - 23.21	6766.7	6766.7 - 6766.7
IL-12	3.36	1.94 - 5.83	5.95	2.97 - 22.29	6500	3250 - 6500
RANTES	7.35	4.60 - 11.70	14.4	14.40 - 14.40	1166.7	1166.7 - 2916.7
Eotaxin	0.21	0.12 - 0.35	1.33	1.33 - 1.33	966.7	966.7 - 966.7
IL-17	11.91	8.26 - 17.20	35.48	17.74 - 53.28	12933.3	12933.3 - 12933.3
MIP-1α	10.51	8.26 - 13.20	11.43	11.43 - 34.29	8333.3	8333.3 - 20833.3
GM-CSF	6.26	4.40 - 8.91	23.8	11.90 - 23.80	11566.7	5783.3 - 17350
MIP-1β	13.76	9.09 - 20.80	29.63	29.63 - 29.63	7200	3600 - 7200
MCP-1	2.45	1.80 - 3.32	5.58	2.79 - 8.37	6100	3050 - 6100
IL-15	6.42	4.09 - 10.10	11.11	11.11 - 27.78	8100	8100 - 20250
IL-5	0.95	0.72-1.26	2.14	1.07 - 3.20	2330	1165 - 2330
IFN-γ	1.69	1.19 - 2.41	4.8	4.80 - 4.80	10500	5250 - 10500
IFN-α	9.05	5.15 - 15.90	17.7	8.85 - 44.24	4300	4300 - 10750
IL-1RA	23.78	11.80 - 47.80	38.41	38.26 - 115.20	28000	14000 - 28000
IL-2	0.95	0.54 - 1.65	3.3	1.65 - 4.94	1200	1200 - 1200
IL-7	16.71	11.90 - 23.50	40.33	40.33 - 40.33	6533.3	3266.7 - 9800
IP-10	0.56	0.38 - 0.81	1.23	1.23 - 1.23	900	450 - 900
IL-2R	11.73	6.95 - 19.80	28.81	28.81 - 28.81	21000	10500 - 21000
MIG	7.95	4.03 - 15.70	29.63	12.57 - 29.63	2400	1200 - 2400
IL-4	5.09	4.25 - 6.10	19.48	19.48 - 19.48	14200	14200 - 42600

Table C.1 – Analytical limits for Luminex human cytokine 25-plex (Life Technologies, LHC0009M). Eight plate runs. All values are in pg/ml.

	Assay	LOD	95%CI	LOQ	IQR	ULOQ	IQR
A	**TGF-β1**	11.75	5.19 - 26.59	29.32	7.33 - 117.26	30019	7480 - 30019
	TGF-β2	4.33	1.56 - 12.05	6.79	6.79 - 27.14	6948	5645.3 - 6948
	TGF-β3	1.02	0.59 - 1.77	4.24	4.24 - 4.24	10863.8	4345.5 - 17382
B	**MMP-13**	4.39	3.25 - 5.93	4.47	4.47 - 4.47	4572.3	4572.3 - 4572.3

Table C.2 – Analytical limits for other Luminex assays. [A] Bio-Plex Pro Human TGF-β 3-plex (Bio-Rad, 171-W4001M); 6 plate runs. **[B]** VersaMap Human MMP-13 (R&D Systems, VMAPH-01); 5 plate runs. All values are in pg/ml.

	Assay	LOD	95%CI	LOQ	IQR	ULOQ	IQR
A	**IL-1α**	0.171	0.14 - 0.22	0.610	0.61 - 2.44	2500	2500 - 2500
	IL-6	0.062	0.02 - 0.17	0.610	0.38 - 2.44	2500	2500 - 2500
	IL-8	0.128	0.04 - 0.45	0.610	0.38 - 1.53	2500	2500 - 2500
	TNF-α	0.363	0.17 - 0.76	0.610	0.61 - 2.44	2500	2500 - 2500
B	**MMP-1**	6.22	2.82 - 13.73	24.41	24.41 - 97.66	100000	25000 - 100000
	MMP-3	2.50	0.79 - 7.88	24.41	24.41 - 24.41	100000	81250 - 100000
	MMP-9	89.82	36.42 - 7.88	488.28	122.07 - 488.28	500000	406250 - 500000
C	**TIMP-1**	20.75	16.64 - 26.04	54.87	54.87 - 54.87	40000	4000 - 4000
D	**LIGHT**	0.723	0.31 - 1.69	2.440	1.43 - 6.10	2500	2500 - 2500
	DCR3	20.47	4.34 - 96.64	39.060	26.92 - 144.79	10000	10000 - 10000
	BMP-2	25.29	17.96 - 35.60	39.060	39.06 - 42.32	10000	10000 - 10000
	BMP-7	1.66	0.12 - 22.36	14.810	8.79 - 31.25	2000	2000 - 2000
E	**ARGS**	0.29	0.08 - 1.05	0.98	0.98 - 1.95	1000	1000 - 1000

Table C.3 – Analytical limits for Meso Scale Discovery assays. [A] Human pro-inflammatory II 4-plex assay: ultra-sensitive kit (K15025C), 5 plate runs **[B]** Human MMP 3-plex (K15034A), 6 plate runs. **[C]** Human TIMP-1 (K151JFC), 4 plate runs. **[D]** Prototype 4-plex (N45ZA-1), 5 plate runs. **[E]** 374-ARGS neo-epitope (custom GSK in-house), 5 plate runs. All values are in pg/ml except ARGs, which is in ng/ml.

	Assay	LOD	95%CI	LOQ	IQR	ULOQ	IQR
A	**ADAMTS-4**	5.53	1.33 - 22.91	15.63	15.63 - 27.34	312.5	125 - 875
B	**COMP**	0.11	0.03 - 0.41	4.00	4 - 4	128	128 - 128
C	**PIIANP**	39.57	25.09 - 62.42	164.07	65.63 - 459.38	2100	2100 - 2100

Table C.4 – Analytical limits for ELISAs. [A] Human aggrecanase ADAMTS-4 (CusaBio, CSB-EL0001311HU) **[B]** Human COMP (BioVendor, RD194080200), 4 plate runs. **[C]** Human PIIANP (Millipore, EZIIANP-53K), 4 plate runs. All values are in ng/ml.

Appendix D

Quality Control

Quality control data is provided for each patient cohort in Part II of the thesis.

- The proportion of samples above limit of detection (LOD) and limit of quantification (LOQ) for each marker.
- The proportion of *detectable samples* with a replicate coefficient of variation (CV) <20% and < 25% is given as a measure of intra-assay precision. The replicate CV is only given if the marker was detectable in at least 50% of samples.

Marker	>LOD		>LOQ		CV<20%		CV<25%	
IL-1β	86.7%	13/15	46.7%	7/15	46.7%	7/13	46.7%	7/13
TNF-α	100%	15/15	93.3%	14/15	73.3%	11/15	73.3%	11/15
IL-6	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IL-8	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IL-2	21.4%	3/14	0%	0/14				
IL-12	100%	14/14	100%	14/14	85.7%	12/14	85.7%	12/14
IL-15	100%	14/14	100%	14/14	78.6%	11/14	85.7%	12/14
GM-CSF	7.1%	1/14	0%	0/14				
IL-1Ra	71.4%	10/14	21.4%	3/14	60%	6/10	70%	7/10
IL-4	0%	0/14	0%	0/14				
IL-10	14.3%	2/14	0%	0/14				
IL-2R	100%	14/14	21.4%	3/14	42.9%	6/14	64.3%	9/14
RANTES	42.9%	6/14	35.7%	5/14				
MIP-1α	21.4%	3/14	21.4%	3/14				
MIP-1β	35.7%	4/14	7.1%	1/14				
MCP-1	100%	14/14	100%	14/14	71.4%	10/14	71.4%	10/14
IP-10	100%	14/14	100%	14/14	85.7%	12/14	85.7%	12/14
Eotaxin	100%	14/14	100%	14/14	64.3%	9/14	64.3%	9/14
MIG	21.4%	3/14	0%	0/14				
TGF-β1	100%	15/15	100%	15/15	93.3%	15/14	93.3%	14/15
TGF-β2	100%	15/15	100%	15/15	86.7%	13/15	93.3%	14/15
TGF-β3	100%	15/15	100%	15/15	93.3%	14/15	93.3%	14/15
BMP-2	46.7%	7/15	13.3%	2/15				
BMP-7	0%	0/15	0%	0/15				
MMP-1	100%	15/15	100%	15/15	100%	15/15	100%	15/15
MMP-3	100%	15/15	100%	15/15	100%	15/15	100%	15/15
MMP-9	100%	15/15	93.3%	14/15	66.7%	10/15	73.3%	11/15
MMP-13	46.7%	7/15	40%	6/15	85.7%	6/7	100%	7/7
TIMP-1	100%	15/15	100%	15/15	100%	15/15	100%	15/15
ADAMTS-4	100%	15/15	100%	15/15	100%	15/15	100%	15/15
ARGS	93.3%	14/15	93.3%	14/15	92.7%	13/14	100%	14/14
COMP	100%	15/15	100%	15/15	100%	15/15	100%	15/15
PIIANP	100%	15/15	86.7%	13/15	85.7%	12/15	85.7%	12/15
LIGHT	0%	0/15	0%	0/15				
DcR3	100%	15/15	80%	12/15	100%	15/15	100%	15/15

Table D.1 – Isolated anteromedial knee OA (AMOA). N=15

Marker	>LOD		>LOQ		CV<20%		CV<25%	
IL-1β	93.3%	14/15	53.3%	8/15	64.3%	9/14	71.4%	10/14
TNF-α	86.7%	13/15	86.6%	13/15	80%	12.13	86.7%	13/15
IL-6	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IL-8	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IL-2	0%	0/14	0%	0/14				
IL-12	100%	14/14	100%	14/14	100%	14/14	100%	14/14
IL-15	92.9%	13/15	92.9%	13/14	100%	13/13	100%	13/13
GM-CSF	21.4%	3/14	0%	0/14				
IL-1Ra	57.1%	8/14	42.9%	6/14	50%	4/8	62.5%	5/8
IL-4	0%	0/14	0%	0/14				
IL-10	7.1%	1/14	0%	0/14				
IL-2R	100%	14/14	42.9%	6/14	71.4%	10/14	85.7%	12/14
RANTES	35.7%	5/14	28.6%	4/14				
MIP-1α	50%	7/14	35.7%	5/14	100%	7/7	100%	7/7
MIP-1β	42.9%	6/14	7.1%	1/14				
MCP-1	100%	14/14	100%	14/14	100%	14/14	100%	14/14
IP-10	100%	14/14	100%	14/14	100%	14/14	100%	14/14
Eotaxin	100%	14/14	100%	14/14	100%	14/14	100%	14/14
MIG	35.7%	5/14	14.3%	2/14				
TGF-β1	100%	15/15	100%	15/15	100%	15/15	100%	15/15
TGF-β2	100%	15/15	100%	15/15	100%	15/15	100%	15/15
TGF-β3	100%	15/15	100%	15/15	100%	15/15	100%	15/15
BMP-2	64.3%	9/14	35.7%	5/14	100%	9/9	100%	9/9
BMP-7	7.1%	1/14	0%	0/14				
MMP-1	100%	14/14	100%	14/14	100%	14/14	100%	14/14
MMP-3	100%	14/14	100%	14/14	100%	14/14	100%	14/14
MMP-9	100%	14/14	85.7%	12/14	78.6%	11/14	85.7%	12/14
MMP-13	57.1%	8/14	57.1%	8/14	80%	8/10	80%	8/10
TIMP-1	100%	14/14	100%	14/14	100%	14/14	100%	14/14
ADAMTS-4	100%	14/14	100%	14/14	85.7%	12/14	92.9%	13/14
ARGS	100%	14/14	100%	14/14	100%	14/14	100%	14/14
COMP	100%	14/14	100%	14/14	100%	14/14	100%	14/14
PIIANP	100%	14/14	100%	14/14	92.9%	13/14	100%	14/14
LIGHT	28.6%	4/14	7.1%	1/14				
DcR3	85.7%	12/14	71.4%	10/14	91.7%	11/12	100%	12/12

Table D.2 – Anteromedial knee OA with patello-femoral OA (AMOAp). N=15

Marker	>LOD		>LOQ		CV<20%		CV<25%	
IL-1β	90%	27/30	30%	9/30	55.6%	15/27	55.6%	15/27
TNF-α	93.3%	28/30	86.7%	26/30	85.7%	26/28	85.7%	26/28
IL-6	100%	30/30	100%	30/30	100%	30/30	100%	30/30
IL-8	100%	30/30	100%	30/30	100%	30/30	100%	30/30
IL-2	6.9%	2/29	0%	0/29				
IL-12	100%	29/29	100%	29/29	82.8%	24/29	89.7%	26/29
IL-15	100%	29/29	96.6%	28/29	72.4%	21/29	82.8%	24/29
GM-CSF	20.7%	6/29	0%	0/29				
IL-1Ra	65.5%	19/29	31.0%	9/29	68.4%	13/19	68.4%	13/19
IL-4	3.4%	1/29	3.4%	1/29				
IL-10	10.3%	3/29	0%	0/29				
IL-2R	89.7%	26/29	34.5%	10/29	73.1%	19/26	80.8%	21/26
RANTES	34.5%	20/29	17.2%	5/29				
MIP-1α	44.8%	13/29	44.8%	13/29				
MIP-1β	31.0%	9/29	3.4%	1/29				
MCP-1	100%	29/29	100%	29/29	72.4%	21/29	82.8%	24/29
IP-10	100%	29/29	100%	29/29	89.7%	26/29	89.7%	26/29
Eotaxin	100%	29/29	89.7%	26/29	72.4%	21/29	75.9%	22/29
MIG	37.9%	11/29	13.8%	4/29				
TGF-β1	100%	30/30	100%	30/30	96.7%	29/30	96.7%	29/30
TGF-β2	100%	30/30	100%	30/30	96.7%	29/30	96.7%	29/30
TGF-β3	100%	30/30	100%	30/30	96.7%	29/30	96.7%	29/30
BMP-2	69%	20/29	17.2%	5/29	100%	20/20	100%	20/20
BMP-7	17.2%	5/29	0%	0/29				
MMP-1	100%	29/29	100%	29/29	96.6%	28/29	100%	29/29
MMP-3	100%	29/29	100%	29/29	100%	29/29	100%	29/29
MMP-9	100%	29/29	79.3%	23/29	72.4%	21/29	89.7%	26/29
MMP-13	48.3%	14/29	44.8%	13/29				
TIMP-1	100%	29/29	100%	29/29	100%	29/29	100%	29/29
ADAMTS-4	100%	29/29	100%	29/29	96.6%	28/29	100%	29/29
ARGS	93.1%	27/29	93.1%	27/29	96.3%	26/29	96.3%	26/29
COMP	100%	29/29	100%	29/29	100%	29/29	100%	29/29
PIIANP	100%	29/29	82.1%	23/28	89.3%	25/29	92.9%	26/29
LIGHT	13.8%	4/19	0%	0/29				
DcR3	96.6%	28/29	82.8%	24/29	96.4%	27/28	100%	28/28

Table D.3 – Tri-compartmental knee OA. N=30

Marker	>LOD		>LOQ		CV<20%		CV<25%	
IL-1β	63.2%	12/19	10.5%	2/19	91.7%	11/12	100%	12/12
TNF-α	52.6%	10/19	52.6%	10/19	52.6%	10/10	100%	10/10
IL-6	78.9%	15/19	63.2%	12/19	86.7%	13/15	86.7%	13/15
IL-8	100%	19/19	100%	19/19	94.7%	18/19	94.7%	18/19
IL-2	0%	0/19	0%	0/19				
IL-12	100%	19/19	89.5%	17/19	89.5%	17/19	100%	19/19
IL-15	100%	19/19	94.7%	18/19	84.2%	16/19	89.5%	17/19
GM-CSF	21.1%	4/19	5.3%	1/19				
IL-1Ra	0%	0/19	0%	0/19				
IL-4	0%	0/19	0%	0/19				
IL-10	5.3%	1/19	0%	0/19				
IL-2R	26.3%	5/19	0%	0/19				
RANTES	42.1%	8/19	36.8%	7/19				
MIP-1α	0%	0/19	0%	0/19				
MIP-1β	5.3%	1/19	0%	0/19				
MCP-1	100%	19/19	100%	19/19	63.2%	12/19	68.4%	13/19
IP-10	100%	19/19	100%	19/19	100%	19/19	100%	19/19
Eotaxin	100%	19/19	94.7%	18/19	63.2%	12/19	63.2%	12/19
MIG	5.3%	1/19	0%	0/19				
TGF-β1	100%	20/20	100%	20/20	100%	20/20	100%	20/20
TGF-β2	100%	20/20	100%	20/20	100%	20/20	100%	20/20
TGF-β3	100%	20/20	100%	20/20	100%	20/20	100%	20/20
BMP-2	41.2%	7/17	0%	0/17				
BMP-7	11.8%	2/17	0%	0/17				
MMP-1	100%	16/16	93.8%	15/16	100%	16/16	100%	16/16
MMP-3	100%	16/16	100%	16/16	100%	16/16	100%	16/16
MMP-9	100%	16/16	68.8%	11/16	68.8	11/16	75%	12/16
MMP-13	5%	1/20	5.0%	1/20				
TIMP-1	100%	16/16	93.8%	15/16	100%	16/16	100%	16/16
ADAMTS-4	100%	15/15	100%	15/15	100%	15/15	100%	15/15
ARGS	100%	16/16	100%	16/16	100%	16/16	100%	16/16
COMP	100%	16/16	100%	16/16	100%	16/16	100%	16/16
PIIANP	100%	16/16	64.7%	11/17	93.4%	15/16	93.4%	15/16
LIGHT	17.6%	3/17	0%	0/17				
DcR3	94.4%	16/17	64.7%	11/17	82.4%	14/17	88.2%	15/17

Table D.4 – Knee injury. N=20

Marker	>LOD		>LOQ		CV<20%		CV<25%	
IL-1β	100%	18/18	100%	18/18	83.3%	15/18	83.3%	15/18
TNF-α	100%	18/18	100%	18/18	100%	18/18	100%	18/18
IL-6	100%	18/18	100%	18/18	100%	18/18	100%	18/18
IL-8	100%	18/18	100%	18/18	100%	18/18	100%	18/18
IL-2	33.3%	5/15	6.7%	1/15				
IL-12	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IL-15	100%	15/15	100%	15/15	93.3%	14/15	93/3%	14/15
GM-CSF	66.7%	10/15	53.3%	8/15	100%	10/10	100%	10/10
IL-1Ra	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IL-4	93.3%	14/15	13.3%	2/15	100%	14/14	100%	14/14
IL-10	93.3%	14/15	46.7	7/15	100%	14/14	100%	14/14
IL-2R	93.3%	14/15	93.3%	14/15	100%	14/14	100%	14/14
RANTES	93.3%	14/15	80%	12/15	92.9%	13/14	92.9%	13/14
MIP-1α	100%	15/15	86.7%	13/15	100%	15/15	100%	15/15
MIP-1β	73.3%	11/15	46.7%	7/15	100%	11/11	100%	11/11
MCP-1	100%	15/15	100%	15/15	100%	15/15	100%	15/15
IP-10	100%	15/15	100%	15/15	100%	15/15	100%	15/15
Eotaxin	100%	15/15	93.3%	14/15	100%	15/15	100%	15/15
MIG	80%	12/15	60%	9/15	100%	12/12	100%	12/12
TGF-β1	100%	15/15	100%	15/15	100%	15/15	100%	15/15
TGF-β2	100%	15/15	100%	15/15	100%	15/15	100%	15/15
TGF-β3	100%	15/15	100%	15/15	100%	15/15	100%	15/15
BMP-2	66.7%	12/18	55.6%	10/18	100%	12/12	100%	12/12
BMP-7	5.6%	1/18	0%	0/18				
MMP-1	100%	18/18	100%	18/18	100%	18/18	100%	18/18
MMP-3	100%	18/18	100%	18/18	94.4%	17/18	94.4%	17/18
MMP-9	100%	18/18	88.9%	16/18	94.4%	17/18	94.4%	17/18
MMP-13	53.8%	7/13	46.2%	6/18	71.4%	5/7	100%	7/7
TIMP-1	100%	18/18	100%	18/18	100%	18/18	100%	18/18
ADAMTS-4	100%	18/18	100%	18/18	100%	18/18	100%	18/18
ARGS	94.4%	17/18	72.2%	13/18	100%	17/17	100%	17/17
COMP	100%	18/18	88.9%	16/18	100%	18/18	100%	18/18
PIIANP	100%	18/18	44.4%	8/18	94.4%	17/18	100%	18/18
LIGHT	88.9%	16/18	83.3%	15/18	100%	16/16	100%	16/16
DcR3	94.4%	17/18	94.4%	17/18	88.2%	15/17	94.1%	16/17

Table D.5 – Inflammatory arthritis. N=18

Marker	>LOD		>LOQ		CV<20%		CV<25%	
IL-1β	80%	8/10	20%	2/10	25%	2/8	37.5%	3/8
TNF-α	100%	10/10	100%	10/10	90%	9/10	100%	10/10
IL-6	100%	10/10	100%	10/10	100%	10/10	100%	10/10
IL-8	100%	10/10	100%	10/10	100%	10/10	100%	10/10
IL-2	0%	0/10	0%	0/10				
IL-12	100%	10/10	100%	10/10	100%	10/10	100%	10/10
IL-15	100%	10/10	100%	10/10	70%	7/10	80%	8/10
GM-CSF	30%	3/10	10%	1/10				
IL-1Ra	50%	5/10	30%	3/10	60%	3/5	60%	3/5
IL-4	0%	0/10	0%	0/10				
IL-10	10%	1/10	0%	0/10				
IL-2R	70%	7/10	20%	2/10	71.4%	5/7	71.4%	5/7
RANTES	20%	2/10	10%	1/10				
MIP-1α	90%	9/10	70%	7/10	77.8%	7/9	88.9%	8/9
MIP-1β	20%	2/10	0%	0/10				
MCP-1	100%	10/10	100%	10/10	100%	10/10	100%	10/10
IP-10	100%	10/10	100%	10/10	100%	10/10	100%	10/10
Eotaxin	100%	10/10	100%	10/10	100%	10/10	100%	10/10
MIG	20%	2/10	0%	0/10				
TGF-β1	100%	10/10	100%	10/10	100%	10/10	100%	10/10
TGF-β2	100%	10/10	100%	10/10	100%	10/10	100%	10/10
TGF-β3	100%	10/10	100%	10/10	100%	10/10	100%	10/10
BMP-2	80%	8/10	10%	1/10	100%	8/8	100%	8/8
BMP-7	0%	0/10	0%	0/10				
MMP-1	100%	9/9	100%	9/9	100%	9/9	100%	9/9
MMP-3	100%	9/9	100%	9/9	100%	9/9	100%	9/9
MMP-9	100%	9/9	100%	9/9	88.9%	8/9	88.9%	8/9
MMP-13	70%	7/10	70%	7/10	100%	7/7	100%	7/7
TIMP-1	100%	9/9	100%	9/9	100%	9/9	100%	9/9
ADAMTS-4	100%	9/9	100%	9/9	100%	9/9	100%	9/9
ARGS	100%	9/9	100%	9/9	100%	9/9	100%	9/9
COMP	100%	9/9	100%	9/9	100%	9/9	100%	9/9
PIIANP	100%	9/9	77/8%	7/9	100%	9/9	100%	9/9
LIGHT	60%	6/10	0%	0/10	100%	6/6	100%	6/6
DcR3	100%	10/10	100%	10/10	100%	10/10	100%	10/10

Table D.6 – Lateral compartment knee OA (Validation test group). N=10

Appendix E

Reliability of Radiographic Grading

All standardised radiographs in this thesis were evaluated using the Kellgren & Lawrence (KL) and OARSI Atlas grading systems as described in Chapter 2 [Section 2.5.2 on page 66]. A single rater (Chethan Jayadev) conducted all radiographic assessments.

Intra-observer (test - retest) reliability: Twenty-four patients' radiographs were chosen at random and re-assessed approximately 3 months later by the original rater (C.J) [Table E.1]. Patients included 2 knee injury, 12 anteromedial knee OA with or without patello-femoral OA, 9 tri-compartmental OA, and one lateral compartment knee OA.

Inter-observer reliability: Sixteen patients' radiographs were chosen at random and assessed by a second evaluator (Ben Davies) [Table E.2] . Patients included 2 knee injury, 6 anteromedial knee OA with or without patello-femoral OA, 7 tri-compartmental OA, and one lateral compartment knee OA.

The kappa coefficient is a measure of agreement [504]; a value >0 indicates that the agreement observed is greater than that expected by chance and a value of 1.0 indicates perfect agreement. Simple kappa coefficients were used to evaluate agreement between the two readings when the variable was dichotomous. Linear-weighted kappa coefficients were used if the variable had more than two ordinal categories.

Radiographic Parameter	kappa	SE
KL Medial	0.624	0.144
KL Lateral	0.848	0.134
KL PF	0.716	0.132
JSN Medial	0.939	0.167
JSN Lateral	0.784	0.152
OP Medial Tibia	0.768	0.147
OP Medial Femur	0.744	0.148
OP Lateral Tibia	0.871	0.151
OP Lateral Femur	0.604	0.154
Sclerosis Medial Tibia	1.000	0.204
Sclerosis Lateral Femur	0.813	0.204
Attrition Medial Tibia	0.654	0.192

Table E.1 – Intra-observer (test - retest) reliability for radiographic assessments. KL - The Kellgren & Lawrence grade; PF - patello-femoral compartment; JSN - joint space narrowing; OP - osteophyte

Radiographic Parameter	kappa	SE
KL Medial	0.628	0.173
KL Lateral	0.633	0.162
KL PF	0.692	0.164
JSN Medial	0.746	0.176
JSN Lateral	0.837	0.198
OP Medial Tibia	0.576	0.169
OP Medial Femur	0.623	0.168
OP Lateral Tibia	0.82	0.192
OP Lateral Femur	0.509	0.173
Sclerosis Medial Tibia	0.158	0.135
Sclerosis Lateral Femur	0.31	0.181
Attrition Medial Tibia	0.714	0.24

Table E.2 – Inter-observer reliability for radiographic assessments. KL - The Kellgren & Lawrence grade; PF - patello-femoral compartment; JSN - joint space narrowing; OP - osteophyte

Appendix F

Part II Supplementary Data

F.1 PLS-DA Regression Diagnostics

Figures F.1, F.2 and F.3.

The Normal Probability Plot of Residuals

This displays the PLS-regression residuals standardised on a double log scale. The standardised residual is the raw residual divided by the residual standard deviation (RSD). This plot allows detection of outliers and assessment of the normality of the residuals. If the residuals are random and normally distributed, the normal probability plot of the residuals has all the points lying on a straight line between –4 and +4 standardised standard deviations. Experimental runs lying outside the –4 or +4 standard deviations are outliers.

The Permutations Plot

This helps to assess the risk that the current model is spurious, i.e. the model just fits the training set well but does not predict Y well for new observations. The R^2 and Q^2 values of the original model are compared to those from several models generated by the order of the Y-observations which has been randomly permuted, while the X-matrix has been kept intact. The plot shows, for a selected Y-variable, on the vertical axis the

values of R^2 and Q^2 for the original model (far to the right) and of the Y-permuted models further to the left. The horizontal axis shows the correlation between the permuted Y-vectors and the original Y-vector for the selected Y. The original Y has the correlation 1.0 with itself, defining the high point on the horizontal axis. The criteria for validity are:

- All permuted blue Q^2-values to the left are lower than the original points to the right.
- The blue regression line of the Q^2-points intersects the vertical axis (on the left) at, or below zero.
- All green R^2-values to the left are lower than the original point to the right.

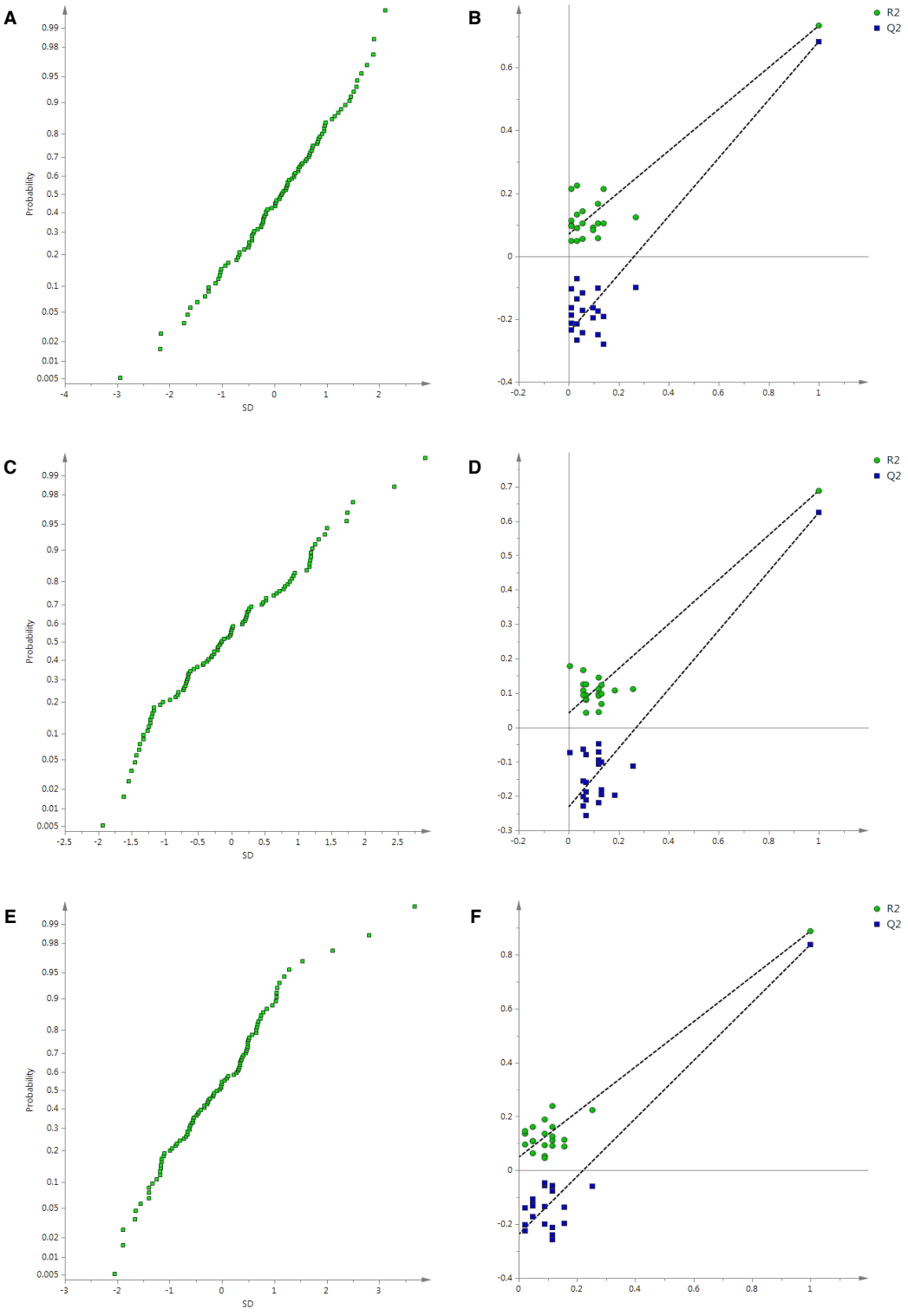


Figure F.1 – Model diagnostics for full PLS-DA model from 20 quantitative markers. Three-component model: R^2=0.770; Q^2=0.718. Normal Probability Plot of Residuals for **[A]** end-stage knee OA, **[C]** knee injury and **[E]** inflammatory arthritis. Permutations Plot (20 permutations) for **[B]** end-stage knee OA, **[D]** knee injury and **[F]** inflammatory arthritis.

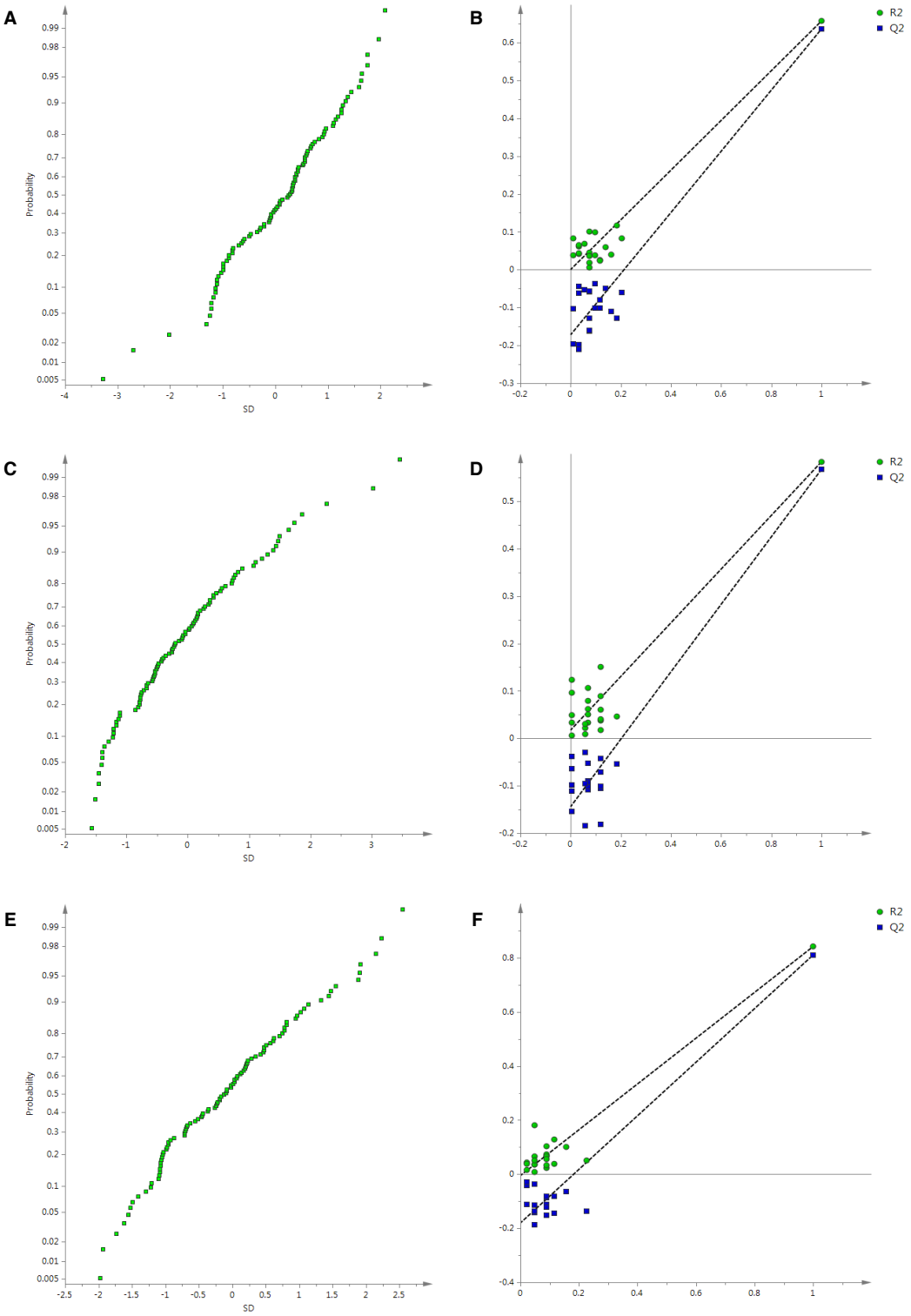


Figure F.2 – Model diagnostics for streamlined quantitative PLS-DA model. Two-component model: R^2=0.694; Q^2=0.673. *Normal Probability Plot of Residuals* for **[A]** end-stage knee OA, **[C]** knee injury and **[E]** inflammatory arthritis. *Permutations Plot* (20 permutations) for **[B]** end-stage knee OA, **[D]** knee injury and **[F]** inflammatory arthritis.

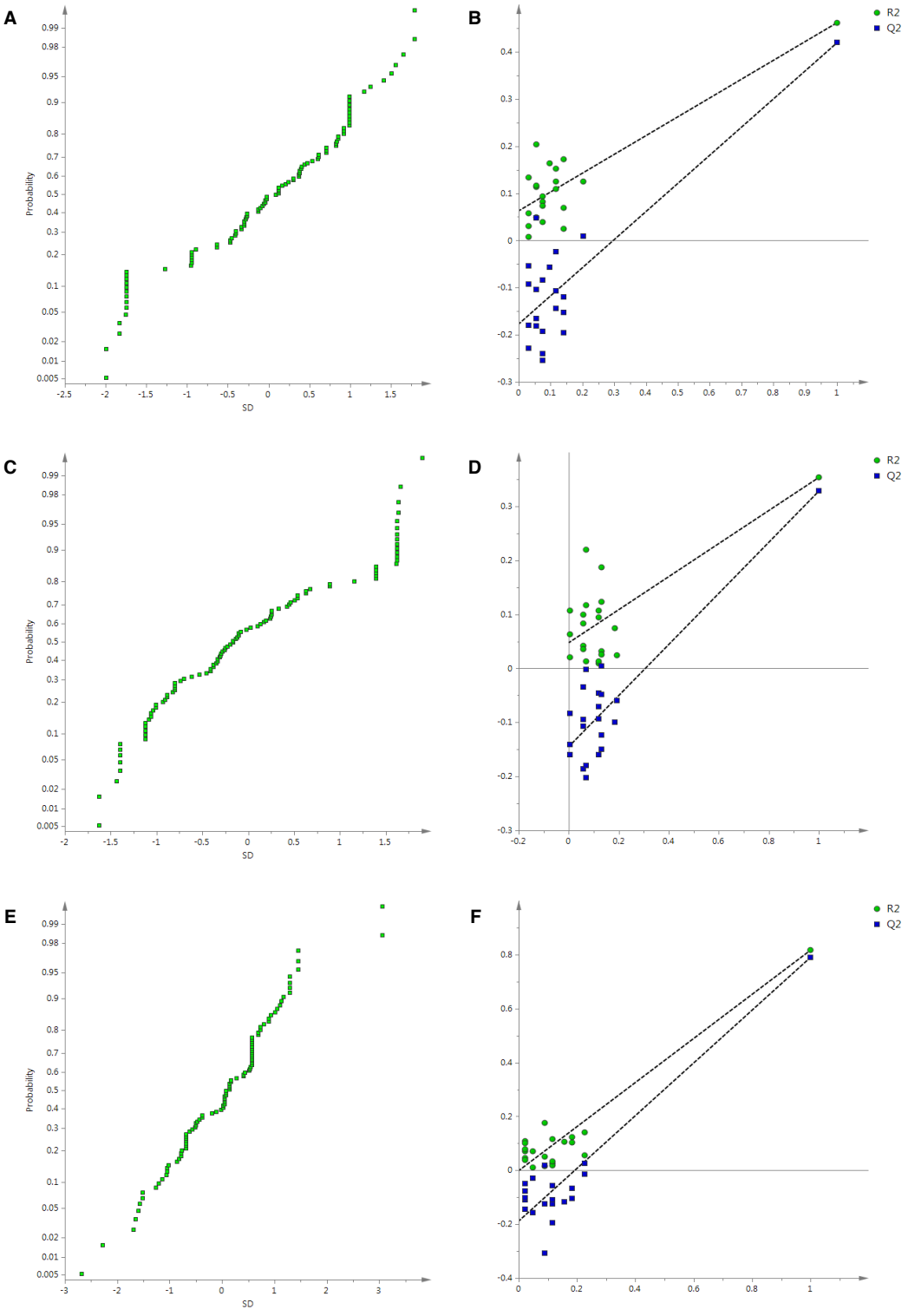


Figure F.3 – Model diagnostics for PLS-DA model from 8 categorical markers. Three-component model: R^2=0.540; Q^2=0.508. *Normal Probability Plot of Residuals* for **[A]** end-stage knee OA, **[C]** knee injury and **[E]** inflammatory arthritis. *Permutations Plot* (20 permutations) for **[B]** end-stage knee OA, **[D]** knee injury and **[F]** inflammatory arthritis.

F.2 Correlation Analysis

There were no important correlations ($\rho > 0.30$ or $\rho < -0.30$) between SF analytes and clinical assessments (OKS, OKS-Pain and OKS-Function) in patients with end-stage knee OA [Figure F.4]. However, there was a suggestion that TGF-β2 and TGF-β3 were associated with better clinical assessments. TGF-β3 showed correlations with better OKS ($\rho = 0.26$, $p = 0.046$) and OKS-Pain ($\rho = 0.27$, $p = 0.038$) scores that fell just below the threshold of interest, but neither remained significant after correction for multiple testing.

There was only one notable correlation between SF analytes and radiographic severity; DcR3 showed a weak positive correlation (Spearman's $\rho = 0.32$, $p = 0.015$) with the OARSI joint space narrowing score, which was lost after correction for multiple testing [Figure F.4]. Pro-inflammatory cytokines (TNF-α, IL-6 and IL-8) appear to be associated with greater radiographic severity in OA. The correlation of IL-8 with worse total KLG ($\rho = 0.29$, $p = 0.024$) fell just below threshold and lost significance after multiple testing correction. MMP-1, MMP-3 and MMP-9 appear to be associated with worse clinical scores but better radiographic scores, especially those weighted with osteophytosis. The correlation of MMP-3 with better total KLG ($\rho = -0.25$, $p = 0.061$) and OARSI total ($\rho = -0.26$, $p = 0.050$) fell below importance and/or lost significance after correction of the family wise error. The aggrecanase, ADAMTS-4 and 374-ARGS neo-epitope both appear to be associated with less radiographic severity, whereas COMP appears to trend with worse scores especially those weighted to osteophytosis. The correlation of ADAMTS-4 with better OARSI JSN ($\rho = -0.27$, $p = 0.038$), and COMP with worse total KLG ($\rho = 0.29$, $p = 0.028$) both fell short of notable and lost significance after correction of the family wise error.

The majority of SF analytes appeared to be associated with worse clinical severity in knee injury patients [Figure F.5]. IL-15 ($\rho = -0.30$) and DcR3 ($\rho = -0.32$) were associated with worse OKS-Function. There were moderate correlations between MMP-1 and worsening overall OKS ($\rho = -0.38$ and $\rho = -0.42$) and OKS-Pain scores ($\rho = -0.34$ and $\rho = -0.50$). IP-10 was moderately correlated with deterioration in all three clinical indices ($\rho \sim -0.4$). COMP on the other hand, showed a moderate correlation with

better OKS ($\rho = -0.38$) and OKS-Pain ($\rho = -0.50$). None of the correlations between SF analytes and clinical scores were significant, even before correction for multiple testing.

The relationship of time of SF sampling from index injury (interval) with both SF measurements and clinical scores was assessed as a potential confounding factor. All three OKS scores improved with time, with OKS ($\rho = 0.33$) and OKS-Pain ($\rho = 0.34$) showing the most notable change. Similarly most SF analytes levels were either relatively unchanged or reduced with time interval. In particular IL-6 ($\rho = -0.36$), TGF-β1 ($\rho = -0.33$), MMP-3 ($\rho = -0.41$) and TIMP-1 ($\rho = -0.33$). Conversely, IL-8 ($\rho = 0.26$), Eotaxin ($\rho = 0.29$) and most markedly IL-12 ($\rho = 0.39$) increase with time from injury. However, all of these correlations were small to moderate at most and none were statistically significant even before correcting for multiple testing.

	OKS	OKS Pain	OKS Func	KLG Total	OARSI Total	OARSI JNS	OARSI OP
TNF-α	-0.01	-0.04	0.01	0.16	0.04	0.11	0.02
IL-6	-0.08	-0.01	-0.16	0.12	0.09	0.08	0.13
IL-8	-0.06	0.00	-0.11	0.29	0.19	0.23	0.19
IL-12	0.05	0.01	0.07	0.00	0.03	0.13	-0.02
IL-15	0.07	0.05	0.07	-0.11	-0.04	0.03	-0.09
DcR-3	-0.12	-0.07	-0.18	0.18	0.21	0.32	0.17
MCP-1	-0.05	0.00	-0.08	-0.02	-0.02	0.04	-0.03
IP-10	0.09	0.08	0.07	-0.12	-0.07	0.04	-0.14
Eotaxin	-0.10	-0.06	-0.13	-0.09	0.02	0.01	0.02
TGF-β1	0.04	0.04	0.04	-0.02	0.08	-0.05	0.09
TGF-β2	0.17	0.18	0.15	0.13	0.04	0.08	0.02
TGF-β3	0.26	0.27	0.21	0.03	-0.01	-0.02	-0.04
MMP-1	-0.23	-0.21	-0.24	-0.19	-0.18	-0.06	-0.17
MMP-3	-0.21	-0.20	-0.20	-0.25	-0.26	-0.07	-0.24
MMP-9	-0.12	-0.13	-0.10	-0.18	-0.23	-0.15	-0.19
MMP-13	-0.01	-0.01	0.02	0.16	0.00	0.07	0.03
ADAMTS-4	0.09	0.07	0.14	-0.12	-0.17	-0.27	-0.13
TIMP-1	-0.13	-0.06	-0.19	0.11	0.10	0.10	0.12
ARGS	-0.11	-0.10	-0.10	-0.21	-0.24	-0.16	-0.21
COMP	0.03	0.05	0.03	0.29	0.22	0.08	0.24
PIIANP	0.11	0.13	0.08	-0.03	-0.01	-0.07	0.00

1.0
-1.0

Figure F.4 – Correlation map for end-stage knee OA. N=60. Spearman's rank correlation coefficients between SF analytes and clinical and radiographic scores. Clinical scores: Oxford Knee Score (OKS), OKS Pain & OKS Function (OKS Func). Radiographic scores: total Kellgren Lawrence grade (KLG Total), OARSI total, OARSI joint space narrowing score (OARSI JSN) & OARSI osteophyte score (OARSI OP). Correlation coefficients weaker than +/- 0.3 were considered not important (grey text). No correlations were significant (after correction by the Holm step-down method).

	OKS	OKS Pain	OKS Func	Interval
TNF-α	-0.03	-0.09	-0.09	-0.06
IL-6	-0.25	-0.26	-0.25	-0.36
IL-8	0.11	0.02	0.06	0.26
IL-12	0.20	0.13	0.17	0.39
IL-15	-0.23	-0.25	-0.30	-0.11
DcR-3	-0.21	-0.23	-0.32	-0.09
MCP-1	-0.17	-0.25	-0.06	0.23
IP-10	-0.40	-0.43	-0.40	-0.23
Eotaxin	-0.01	-0.09	0.27	0.29
TGF-β1	-0.06	0.00	-0.04	-0.33
TGF-β2	-0.13	-0.09	-0.09	-0.11
TGF-β3	0.00	0.03	-0.04	-0.20
MMP-1	-0.38	-0.45	-0.31	-0.18
MMP-3	-0.42	-0.50	-0.25	-0.41
MMP-9	-0.03	-0.09	-0.05	0.14
TIMP-1	-0.17	-0.21	-0.27	-0.30
ADAMTS-4	-0.19	-0.29	-0.06	-0.12
ARGS	-0.16	-0.21	-0.09	-0.02
COMP	0.38	0.50	-0.05	0.02
PIIANP	-0.06	-0.05	-0.23	-0.07
Interval	0.33	0.23	0.34	

1.0
-1.0

Figure F.5 – Correlation map for knee injury. N=20. Spearman's rank correlation coefficients between SF analytes and clinical scores: Oxford Knee Score (OKS), OKS Pain & OKS Function (OKS Func). Correlations are also shown for time, in months, between index knee injury and SF sampling (Interval) with SF analytes and clinical scores. Correlation coefficients weaker than +/- 0.3 are ignored (grey text). No correlations were significant (after correction by the Holm step-down method).

F.3 Knee Injury and Inflammatory Arthritis Patients

A description of the heterogeneous knee injury and inflammatory arthritis patients is provided in Tables F.1 & F.2.

Gender	Age	ACL Injury	Meniscal Injury	Interval (months)
M	18	+	+	7
M	29	+	+	13
M	16		+	20
M	32	+	+	6
F	20	+		5
F	27		+	10
M	60		+	2
M	38		+	4
M	21		+	6
M	37		+	4
M	28	+	+	6
F	22	+	+	7
F	22	+		9
M	36	+	+	11
M	29	+		4
M	21	+	+	7
M	23	+		9
M	31	+		40
M	18		+	5
M	24	+		4

Table F.1 – Knee Injury Cohort (N=20). There were 7 patients with isolated meniscal injury; 6 patients with isolated ACL injury; and 7 patients with ACL injury plus meniscal injury. No patients had radiological or arthroscopic evidence of synovitis, osteochondral defects or cartilage wear. The interval between the index injury and sample collection is given in months.The median interval between injury and surgery was 6.5 months (interquartile range 4 – 9.75 months)

Gender	Age	Diagnosis	Duration	Current Treatment(s)
F	68	RA	10	Mycophenolate mofetil
F	71	RA	41	Prednisolone, Leflunomide
F	24	RA	5	
M	32	RA	6	Sulphasalazine
F	63	RA	25	Leflunomide
F	72	RA	5	Azothioprine
F	32	RA	8	Methotrexate, Sulphasalazine
M	28	RA	5	
F	85	RA	14	Leflunomide, Adalimumab
F	53	RA	15	Methotrexate, Sulphasalazine, Hydroxychloroquine
F	65	RA	9	Methotrexate, Sulphasalazine, Adalimumab
M	64	RA	1	Methotrexate, Sulphasalazine, Hydroxychloroquine
M	65	RA	3	Sulphasalazine
M	56	PsA	0	
F	25	PsA	3	
M	25	PsA	0	
M	20	PsA	0	
F	24	PsA	0	

Table F.2 – Inflammatory arthritis cohort (N=18). There were 13 patients with rheumatoid arthritis (RA) and 5 patients with psoriatic arthritis (PsA). Patients with RA were on a range of treatments. The number of years since first diagnosis (duration) is provided.

Bibliography

[1] Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012 Dec;380(9859):2163–2196.

[2] Mankin HJ, Grodzinsky AJ, Buckwalter JA. Articular Cartilage and Osteoarthritis. In: Einhorn TA, O'Keefe RJ, Buckwalter JA, editors. *Orthopaedic Basic Science: Foundations of Clinical Practice*. Rosemont: American Academy of Orthopaedic Surgeons; 2007. p. 161–174.

[3] Eyre DR, Weis MA, Wu JJ. Articular cartilage collagen: an irreplaceable framework? *European Cells and Materials*. 2006;12:57–63.

[4] Halász K, Kassner A, Mörgelin M, Heinegård D. COMP acts as a catalyst in collagen fibrillogenesis. *The Journal of Biological Chemistry*. 2007 Oct;282(43):31166–31173.

[5] Chen FH, Herndon ME, Patel N, Hecht JT, Tuan RS, Lawler J. Interaction of cartilage oligomeric matrix protein/thrombospondin 5 with aggrecan. *The Journal of Biological Chemistry*. 2007 Aug;282(34):24591–24598.

[6] Pfander D, Gelse K. Hypoxia and osteoarthritis: how chondrocytes survive hypoxic environments. *Current Opinion in Rheumatology*. 2007 Sep;19(5):457–462.

[7] Imhof H, Sulzbacher I, Grampp S, Czerny C, Youssefzadeh S, Kainberger F. Subchondral bone and cartilage disease: a rediscovered functional unit. *Investigative Radiology*. 2000 Oct;35(10):581–588.

[8] Lyons TJ, McClure SF, Stoddart RW, McClure J. The normal human chondro-osseous junctional region: evidence for contact of uncalcified cartilage with subchondral bone and marrow spaces. *BMC Musculoskeletal Disorders*. 2006;7:52.

[9] Smith MD. The normal synovium. *The Open Rheumatology Journal*. 2011;5:100–106.

[10] Smith MD, Barg E, Weedon H, Papengelis V, Smeets T, Tak PP, et al. Microarchitecture and protective mechanisms in synovial tissue from clinically and arthroscopically normal knee joints. *Annals of the Rheumatic Diseases*. 2003 Apr;62(4):303–307.

[11] Fam H, Bryant JT, Kontopoulou M. Rheological properties of synovial fluids. *Biorheology*. 2007;44(2):59–74.

[12] Hui AY, McCarty WJ, Masuda K, Firestein GS, Sah RL. A systems biology approach to synovial joint lubrication in health, injury, and disease. *Wiley Interdisciplinary Reviews Systems Biology and Medicine*. 2012 Jan;4(1):15–37.

[13] Rodeo S, Kawamurra S. Form and Function of the Meniscus. In: Einhorn TA, O'Keefe RJ, Buckwalter JA, editors. *Orthopaedic Basic Science: Foundations of Clinical Practice*. Rosemont: American Academy of Orthopaedic Surgeons; 2007. p. 175–190.

[14] Arden N, Nevitt MC. Osteoarthritis: epidemiology. *Best Practice and Research: Clinical Rheumatology*. 2006 Feb;20(1):3–25.

[15] Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor G, Koch G, et al. Lifetime risk of symptomatic knee osteoarthritis. *Arthritis and Rheumatism*. 2008 Sep;59(9):1207–1213.

[16] ARUK. Arthritis in the UK - Key Facts. Arthritis Research UK: Arthritis Research UK; 2008.

[17] Pereira D, Peleteiro B, Araújo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. *Osteoarthritis and Cartilage*. 2011 Nov;19(11):1270–1285.

[18] Kuettner KE, Goldberg VM. Introduction. In: Kuettner KE, Goldberg VM, editors. *Osteoarthritic Disorders*. Rosemont: Academy of Orthopaedic Surgeons; 1995. p. xxi–xxv.

[19] Herrero-Beaumont G, Roman-Blas J, Castañeda S, Jimenez S. Primary Osteoarthritis No Longer Primary: Three Subsets with Distinct Etiological, Clinical, and Therapeutic Characteristics. *Seminars in Arthritis and Rheumatism*. 2009 Jan;39(2):71–80.

[20] Brandt KD, Dieppe P, Radin EL. Commentary: is it useful to subset "primary" osteoarthritis? A critique based on evidence regarding the etiopathogenesis of osteoarthritis. *Seminars in Arthritis and Rheumatism*. 2009 Oct;39(2):81–95.

[21] Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. *BMC Musculoskeletal Disorders*. 2008;9:132.

[22] Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis and Cartilage*. 2010 Jan;18(1):24–33.

[23] Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. *Arthritis and Rheumatism*. 2003 Aug;48(8):2178–2187.

[24] Roos EM. Joint injury causes knee osteoarthritis in young adults. *Current Opinion in Rheumatology*. 2005 Mar;17(2):195–200.

[25] Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. *The American Journal of Sports Medicine*. 2007 Oct;35(10):1756–1769.

[26] Palmer KT. Occupational activities and osteoarthritis of the knee. *British Medical Bulletin*. 2012 Jun;102:147–170.

[27] Brouwer GM, van Tol AW, Bergink AP, Belo JN, Bernsen RMD, Reijman M, et al. Association between valgus and varus alignment and the development and

progression of radiographic osteoarthritis of the knee. *Arthritis and Rheumatism*. 2007 Apr;56(4):1204–1211.

[28] Janakiramanan N, Teichtahl AJ, Wluka AE, Ding C, Jones G, Davis SR, et al. Static knee alignment is associated with the risk of unicompartmental knee cartilage defects. *Journal of Orthopaedic Research*. 2008 Feb;26(2):225–230.

[29] Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. *Osteoarthritis and Cartilage*. 2004;12 Suppl A:S39–44.

[30] Valdes AM, Spector TD. Genetic epidemiology of hip and knee osteoarthritis. *Nature Reviews Rheumatology*. 2011 Jan;7(1):23–32.

[31] Chaganti RK, Lane NE. Risk factors for incident osteoarthritis of the hip and knee. *Current Reviews in Musculoskeletal Medicine*. 2011 Sep;4(3):99–104.

[32] Neogi T, Zhang Y. Epidemiology of osteoarthritis. *Rheumatic Diseases Clinics of North America*. 2013 Feb;39(1):1–19.

[33] Nuki G. Osteoarthritis: a problem of joint failure. *Zeitschrift für Rheumatologie*. 1999 Jun;58(3):142–147.

[34] Brandt KD, Dieppe P, Radin E. Etiopathogenesis of osteoarthritis. *The Medical Clinics of North America*. 2009 Jan;93(1):1–24– xv.

[35] Aigner T, McKenna L. Molecular pathology and pathobiology of osteoarthritic cartilage. *Cellular and Molecular Life Sciences : CMLS*. 2002 Jan;59(1):5–18.

[36] Aigner T, Sachse A, Gebhard P, Roach H. Osteoarthritis: Pathobiology-targets and ways for therapeutic intervention. *Advanced Drug Delivery Reviews*. 2006 Jan;58(2):128–149.

[37] Goldring MB, Goldring SR. Osteoarthritis. *Journal of Cellular Physiology*. 2007 Dec;213(3):626–634.

[38] Abramson SB, Attur M. Developments in the scientific understanding of osteoarthritis. *Arthritis Research and Therapy*. 2009;11(3):227.

[39] Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: A disease of the joint as an organ. *Arthritis and Rheumatism*. 2012 May;64(6):1697–1707.

[40] Murphy G, Knäuper V, Atkinson S, Butler G, English W, Hutton M, et al. Matrix metalloproteinases in arthritic disease. *Arthritis Research*. 2002;4 Suppl 3:S39–49.

[41] Porter S, Clark IM, Kevorkian L, Edwards DR. The ADAMTS metalloproteinases. *The Biochemical Journal*. 2005 Feb;386(Pt 1):15–27.

[42] Nagase H, Visse R, Murphy G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovascular Research*. 2006;69(3).

[43] Goldring MB, Otero M, Plumb DA, Dragomir C, Favero M, El Hachem K, et al. Roles of inflammatory and anabolic cytokines in cartilage metabolism: signals and multiple effectors converge upon MMP-13 regulation in osteoarthritis. *European Cells and Materials*. 2011;21:202–220.

[44] Kashiwagi M. TIMP-3 Is a Potent Inhibitor of Aggrecanase 1 (ADAM-TS4) and Aggrecanase 2 (ADAM-TS5). *Journal of Biological Chemistry*. 2001 Jan;276(16):12501–12504.

[45] Goldring SR, Goldring MB. The role of cytokines in cartilage matrix degeneration in osteoarthritis. *Clinical Orthopaedics and Related Research*. 2004;(427 SUPPL.):S27–S36.

[46] Fernandes JC, Martel-Pelletier J, Pelletier JP. The role of cytokines in osteoarthritis pathophysiology. *Biorheology*. 2002;39(1-2):237–246.

[47] Kapoor M, Martel-Pelletier J, Lajeunesse D, Pelletier JP, Fahmi H. Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nature Reviews Rheumatology*. 2011 Jan;7(1):33–42.

[48] Borzì RM, Mazzetti I, Marcu KB, Facchini A. Chemokines in Cartilage Degradation. *Clinical Orthopaedics and Related Research*. 2004 Oct;427:S53–S61.

[49] Sandell LJ, Xing X, Franz C, Davies S, Chang LW, Patra D. Exuberant expression of chemokine genes by adult human articular chondrocytes in response to IL-1beta. *Osteoarthritis and Cartilage*. 2008 Dec;16(12):1560–1571.

[50] Borzì RM, Mazzetti I, Cattini L, Uguccioni M, Baggiolini M, Facchini A. Human chondrocytes express functional chemokine receptors and release matrix-degrading enzymes in response to C-X-C and C-C chemokines. *Arthritis and Rheumatism*. 2000 Aug;43(8):1734–1741.

[51] Alaaeddine N, Olee T, Hashimoto S, Creighton-Achermann L, Lotz M. Production of the chemokine RANTES by articular chondrocytes and role in cartilage degradation. *Arthritis and Rheumatism*. 2001 Jul;44(7):1633–1643.

[52] Yuan GH, Masuko-Hongo K, Sakata M, Tsuruha J, Onuma H, Nakamura H, et al. The role of C-C chemokines and their receptors in osteoarthritis. *Arthritis and Rheumatism*. 2001 May;44(5):1056–1070.

[53] Blaney Davidson EN, van der Kraan PM, van den Berg WB. TGF-beta and osteoarthritis. *Osteoarthritis and Cartilage*. 2007 Jun;15(6):597–604.

[54] van der Kraan PM, Blaney Davidson EN, van den Berg WB. Bone morphogenetic proteins and articular cartilage: To serve and protect or a wolf in sheep clothing's? *Osteoarthritis and Cartilage*. 2010 Jun;18(6):735–741.

[55] Lories RJU. Joint homeostasis, restoration, and remodeling in osteoarthritis. *Best Practice and Research: Clinical Rheumatology*. 2008 Apr;22(2):209–220.

[56] van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. *Osteoarthritis and Cartilage*. 2007 Mar;15(3):237–244.

[57] Bush JR, Beier F. TGF-β and osteoarthritis–the good and the bad. *Nature Medicine*. 2013 Jun;19(6):667–669.

[58] Henrotin Y, Kurz B, Aigner T. Oxygen and reactive oxygen species in cartilage degradation: friends or foes? *Osteoarthritis and Cartilage*. 2005 Aug;13(8):643–654.

[59] Abramson SB. Osteoarthritis and nitric oxide. *Osteoarthritis and Cartilage*. 2008 Jan;16 Suppl 2:S15–20.

[60] Andriacchi TP, Mündermann A, Smith RL, Alexander EJ, Dyrby CO, Koo S. A Framework for the in Vivo Pathomechanics of Osteoarthritis at the Knee. *Annals of Biomedical Engineering*. 2004;32(3):447–457.

[61] Heinegård D, Saxne T. The role of the cartilage matrix in osteoarthritis. *Nature Reviews Rheumatology*. 2011 Jan;7(1):50–56.

[62] Goldring MB. Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. *Therapeutic Advances in Musculoskeletal Disease*. 2012 Aug;4(4):269–285.

[63] Roach H, Aigner T, Kouri J. Chondroptosis: A variant of apoptotic cell death in chondrocytes? *Apoptosis*. 2004 Jan;9(3):265–277.

[64] del Carlo Jr M, Loeser R. Cell death in osteoarthritis. *Current Rheumatology Reports*. 2008 Jan;10(1):37–42.

[65] Doré S, Pelletier JP, Dibattista JA, Tardif G, Brazeau P, Martel-Pelletier J. Human Osteoarthritic Chondrocytes Possess an Increased Number of Insulin-Like Growth Factor 1 Binding Sites but are Unresponsive to its Stimulation. *Arthritis and Rheumatism*. 1994 Feb;37(2):253–263.

[66] Lories RJ, Luyten FP. The bone-cartilage unit in osteoarthritis. *Nature Reviews Rheumatology*. 2011 Jan;7(1):43–49.

[67] Zhang LZ, Zheng HA, Jiang Y, Tu YH, Jiang PH, Yang AL. Mechanical and biologic link between cartilage and subchondral bone in osteoarthritis. *Arthritis Care & Research*. 2012 Jul;64(7):960–967.

[68] Radin EL, Rose RM. Role of subchondral bone in the initiation and progression of cartilage damage. *Clinical Orthopaedics and Related Research*. 1986 Dec;(213):34–40.

[69] Felson DT, Neogi T. Osteoarthritis: is it a disease of cartilage or of bone? *Arthritis and Rheumatism*. 2004 Feb;50(2):341–344.

[70] Walsh DA, Bonnet CS, Turner EL, Wilson D, Situ M, McWilliams DF. Angiogenesis in the synovium and at the osteochondral junction in osteoarthritis. *Osteoarthritis and Cartilage*. 2007 Jul;15(7):743–751.

[71] Walsh DA, McWilliams DF, Turley MJ, Dixon MR, Fransès RE, Mapp PI, et al. Angiogenesis and nerve growth factor at the osteochondral junction in

rheumatoid arthritis and osteoarthritis. *Rheumatology*. 2010 Oct;49(10):1852–1861.

[72] Burr DB, Gallant MA. Bone remodelling in osteoarthritis. *Nature Reviews Rheumatology*. 2012 Nov;8(11):665–673.

[73] Dürr HD, Martin H, Pellengahr C, Schlemmer M, Maier M, Jansson V. The cause of subchondral bone cysts in osteoarthrosis: a finite element analysis. *Acta Orthopaedica Scandinavica*. 2004 Oct;75(5):554–558.

[74] Daheshia M, Yao JQ. The bone marrow lesion in osteoarthritis. *Rheumatology International*. 2011 Feb;31(2):143–148.

[75] Sellam J, Berenbaum F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nature Reviews Rheumatology*. 2010 Oct;6(11):625–635.

[76] Wenham CYJ, Conaghan PG. The role of synovitis in osteoarthritis. *Therapeutic Advances in Musculoskeletal Disease*. 2010 Dec;2(6):349–359.

[77] Scanzello CR, Goldring SR. The role of synovitis in osteoarthritis pathogenesis. *Bone*. 2012 Aug;51(2):249–257.

[78] de Lange-Brokaar BJE, Ioan-Facsinay A, van Osch GJVM, Zuurmond AM, Schoones J, Toes REM, et al. Synovial inflammation, immune cells and their cytokines in osteoarthritis: a review. *Osteoarthritis and Cartilage*. 2012 Dec;20(12):1484–1499.

[79] Smith MD, Triantafillou S, Parker A, Youssef PP, Coleman M. Synovial membrane inflammation and cytokine production in patients with early osteoarthritis. *The Journal of Rheumatology*. 1997 Feb;24(2):365–371.

[80] Krasnokutsky S, Belitskaya-Lévy I, Bencardino J, Samuels J, Attur M, Regatte R, et al. Quantitative magnetic resonance imaging evidence of synovial proliferation is associated with radiographic severity of knee osteoarthritis. *Arthritis and Rheumatism*. 2011 Oct;63(10):2983–2991.

[81] Benito MJ, Veale DJ, FitzGerald O, van den Berg WB, Bresnihan B. Synovial tissue inflammation in early and late osteoarthritis. *Annals of the Rheumatic Diseases*. 2005 Sep;64(9):1263–1267.

[82] Roemer FW, Guermazi A, Felson DT, Niu J, Nevitt MC, Crema MD, et al. Presence of MRI-detected joint effusion and synovitis increases the risk of cartilage loss in knees without osteoarthritis at 30-month follow-up: the MOST study. *Annals of the Rheumatic Diseases*. 2011 Oct;70(10):1804–1809.

[83] Ayral X, Pickering EH, Woodworth TG, Mackillop N, Dougados M. Synovitis: a potential predictive factor of structural progression of medial tibiofemoral knee osteoarthritis – results of a 1 year longitudinal arthroscopic study in 422 patients. *Osteoarthritis and Cartilage*. 2005 May;13(5):361–367.

[84] Torres L, Dunlop DD, Peterfy C, Guermazi A, Prasad P, Hayes KW, et al. The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. *Osteoarthritis and Cartilage*. 2006 Oct;14(10):1033–1040.

[85] Hill CL, Hunter DJ, Niu J, Clancy M, Guermazi A, Genant H, et al. Synovitis detected on magnetic resonance imaging and its relation to pain and cartilage loss in knee osteoarthritis. *Annals of the Rheumatic Diseases*. 2007 Dec;66(12):1599–1603.

[86] Baker K, Grainger A, Niu J, Clancy M, Guermazi A, Crema M, et al. Relation of synovitis to knee pain using contrast-enhanced MRIs. *Annals of the Rheumatic Diseases*. 2010 Oct;69(10):1779–1783.

[87] Carr AJ. Beyond disability: measuring the social and personal consequences of osteoarthritis. *Osteoarthritis and Cartilage*. 1999 Mar;7(2):230–238.

[88] Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, et al. Understanding the pain experience in hip and knee osteoarthritis - an OARSI/OMERACT initiative. *Osteoarthritis and Cartilage*. 2008;16(4):415–422.

[89] Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee

of the American Rheumatism Association. *Arthritis and Rheumatism*. 1986 Aug;29(8):1039–1049.

[90] Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Annals of the Rheumatic Diseases*. 2010 Mar;69(3):483–489.

[91] Cibere J. Do we need radiographs to diagnose osteoarthritis? *Best Practice and Research: Clinical Rheumatology*. 2006 Feb;20(1):27–38.

[92] Wu CW, Morrell MR, Heinze E, Concoff AL, Wollaston SJ, Arnold EL, et al. Validation of American College of Rheumatology classification criteria for knee osteoarthritis using arthroscopically defined cartilage damage scores. *Seminars in Arthritis and Rheumatism*. 2005 Dec;35(3):197–201.

[93] Hawker GA, Wright JG, Badley EM, Coyte PC. Perceptions of, and willingness to consider, total joint arthroplasty in a population-based cohort of individuals with disabling hip and knee arthritis. *Arthritis and Rheumatism*. 2004 Aug;51(4):635–641.

[94] Gossec L, Hawker G, Davis AM, Maillefert JF, Lohmander LS, Altman R, et al. OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. *The Journal of Rheumatology*. 2007 Jun;34(6):1432–1435.

[95] Gossec L, Paternotte S, Maillefert JF, Combescure C, Conaghan PG, Davis AM, et al. The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. *Osteoarthritis and Cartilage*. 2011 Feb;19(2):147–154.

[96] Hutchings A, Calloway M, Choy E, Hooper M, Hunter DJ, Jordan JM, et al. The Longitudinal Examination of Arthritis Pain (LEAP) study: relationships between weekly fluctuations in patient-rated joint pain and other health outcomes. *The Journal of Rheumatology*. 2007 Nov;34(11):2291–2300.

[97] Gooberman-Hill R, Woolhead G, Mackichan F, Ayis S, Williams S, Dieppe P. Assessing chronic joint pain: lessons from a focus group study. *Arthritis and Rheumatism*. 2007 May;57(4):666–671.

[98] Neogi T, Nevitt MC, Yang M, Curtis JR, Torner J, Felson DT. Consistency of knee pain: Correlates and association with function. *Osteoarthritis and Cartilage*. 2010;18(10):1250–1255.

[99] Lee AS, Ellman MB, Yan D, Kroin JS, Cole BJ, van Wijnen AJ, et al. A current review of molecular mechanisms regarding osteoarthritis and pain. *Gene*. 2013 Sep;527(2):440–447.

[100] van Dijk GM, Dekker J, Veenhof C, van den Ende CHM, Carpa Study Group. Course of functional status and pain in osteoarthritis of the hip or knee: a systematic review of the literature. *Arthritis and Rheumatism*. 2006 Oct;55(5):779–785.

[101] Wise BL, Niu J, Zhang Y, Wang N, Jordan JM, Choy E, et al. Psychological factors and their relation to osteoarthritis pain. *Osteoarthritis and Cartilage*. 2010 Jul;18(7):883–887.

[102] Creamer P, Lethbridge-Cejku M, Hochberg MC. Determinants of pain severity in knee osteoarthritis: effect of demographic and psychosocial variables using 3 pain measures. *The Journal of Rheumatology*. 1999 Aug;26(8):1785–1792.

[103] Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. *Lancet*. 2005 Jan;365(9463):965–973.

[104] Gwilym SE, Pollard TCB, Carr AJ. Understanding pain in osteoarthritis. *Journal of Bone and Joint Surgery (British)*. 2008;90(3):280–287.

[105] Sofat N, Ejindu V, Kiely P. What makes osteoarthritis painful? The evidence for local and central pain processing. *Rheumatology*. 2011 Dec;50(12):2157–2165.

[106] Garratt AM, Brealey S, Gillespie WJ, DAMASK Trial Team. Patient-assessed health instruments for the knee: a structured review. *Rheumatology*. 2004 Nov;43(11):1414–1423.

[107] Beard DJ, Knezevic K, Al-Ali S, Dawson J, Price AJ. The use of outcome measures relating to the knee. *Orthopaedics and Trauma*. 2010;24(4):309–316.

[108] McCormack HM, Horne DJ, Sheather S. Clinical applications of visual analogue scales: a critical review. *Psychological Medicine*. 1988;.

[109] Hawker GA, Davis AM, French MR, Cibere J, Jordan JM, March L, et al. Development and preliminary psychometric testing of a new OA pain measure–an OARSI/OMERACT initiative. *Osteoarthritis and Cartilage*. 2008 Apr;16(4):409–414.

[110] Bellamy N, Buchanan W, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *Journal of Rheumatology*. 1988;15(12):1833–1840.

[111] Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS) - Development of a self-administered outcome measure. *Journal of Orthopaedic and Sports Physical Therapy*. 1998;28(2):88–96.

[112] Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone and Joint Surgery (British)*. 1998 Jan;80(1):63–69.

[113] Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. *Journal of Bone and Joint Surgery (British)*. 2007 Aug;89(8):1010–1014.

[114] Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Annals of the Rheumatic Diseases*. 1957 Dec;16(4):494–502.

[115] Hart DJ, Spector TD. Kellgren & Lawrence grade 1 osteophytes in the knee–doubtful or definite? *Osteoarthritis and Cartilage*. 2003 Feb;11(2):149–150.

[116] Schiphof D, Boers M, Bierma-Zeinstra SMA. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. *Annals of the Rheumatic Diseases*. 2007 Oct;67(7):1034–1036.

[117] Schiphof D, de Klerk BM, Kerkhof HJM, Hofman A, Koes BW, Boers M, et al. Impact of different descriptions of the Kellgren and Lawrence classification criteria on the diagnosis of knee osteoarthritis. *Annals of the Rheumatic Diseases*. 2011 Aug;70(8):1422–1427.

[118] Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. *Annals of the Rheumatic Diseases*. 2011 Nov;70(11):1884–1886.

[119] Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis and Cartilage*. 2007 Jan;15 Suppl A:A1–56.

[120] McAlindon TE, Snow S, Cooper C, Dieppe PA. Radiographic patterns of osteoarthritis of the knee joint in the community: the importance of the patellofemoral joint. *Annals of the Rheumatic Diseases*. 1992 Jul;51(7):844–849.

[121] Davies AP, Vince AS, Shepstone L, Donell ST, Glasgow MM. The radiologic prevalence of patellofemoral osteoarthritis. *Clinical Orthopaedics and Related Research*. 2002 Sep;(402):206–212.

[122] Hinman RS, Crossley KM. Patellofemoral joint osteoarthritis: an important subgroup of knee osteoarthritis. *Rheumatology*. 2007 May;46(7):1057–1062.

[123] Guermazi A, Hayashi D, Eckstein F, Hunter DJ, Duryea J, Roemer FW. Imaging of Osteoarthritis. *Rheumatic Disease Clinics of North America*. 2013 Feb;39(1):67–105.

[124] Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. *Osteoarthritis and Cartilage*. 2008 Nov;16(11):1300–1306.

[125] Marijnissen ACA, Vincken KL, Vos PAJM, Saris DBF, Viergever MA, Bijlsma JWJ, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. *Osteoarthritis and Cartilage*. 2008 Feb;16(2):234–243.

[126] Shamir L, Ling SM, Scott WW, Bos A, Orlov N, Macura TJ, et al. Knee x-ray image analysis method for automated detection of osteoarthritis. *IEEE Transactions on Biomedical Engineering*. 2009 Feb;56(2):407–415.

[127] Piperno M, Hellio-Le Graverand MP, Conrozier T, Bochu M, Mathieu P, Vignon E. Quantitative evaluation of joint space width in femorotibial osteoarthritis: comparison of three radiographic views. *Osteoarthritis and Cartilage*. 1998 Jul;6(4):252–259.

[128] Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, et al. Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. *Skeletal Radiology*. 2003 Mar;32(3):128–132.

[129] Vignon E, Piperno M, Le Graverand MPH, Mazzuca SA, Brandt KD, Mathieu P, et al. Measurement of radiographic joint space width in the tibiofemoral compartment of the osteoarthritic knee: comparison of standing anteroposterior and Lyon schuss views. *Arthritis and Rheumatism*. 2003 Feb;48(2):378–384.

[130] Mazzuca SA, Brandt KD, Buckwalter KA, Lequesne M. Pitfalls in the accurate measurement of joint space narrowing in semiflexed, anteroposterior radiographic imaging of the knee. *Arthritis and Rheumatism*. 2004 Aug;50(8):2508–2515.

[131] Emrani PS, Katz JN, Kessler CL, Reichmann WM, Wright EA, McAlindon TE, et al. Joint space narrowing and Kellgren-Lawrence progression in knee osteoarthritis: an analytic literature synthesis. *Osteoarthritis and Cartilage*. 2008 Aug;16(8):873–882.

[132] Hunter DJ, Zhang YQ, Tu X, Lavalley M, Niu JB, Amin S, et al. Change in joint space width: hyaline articular cartilage loss or alteration in meniscus? *Arthritis and Rheumatism*. 2006 Aug;54(8):2488–2495.

[133] Guermazi A, Roemer FW, Haugen IK, Crema MD, Hayashi D. MRI-based semiquantitative scoring of joint pathology in osteoarthritis. *Nature Reviews Rheumatology*. 2013 Apr;9(4):236–251.

[134] Wang Y, Wluka AE, Jones G, Ding C, Cicuttini FM. Use magnetic resonance imaging to assess articular cartilage. *Therapeutic Advances in Musculoskeletal Disease*. 2012 Mar;4(2):77–97.

[135] Guermazi A, Roemer FW, Hayashi D, Crema MD, Niu J, Zhang Y, et al. Assessment of synovitis with contrast-enhanced MRI using a whole-joint semiquantitative scoring system in people with, or at high risk of, knee osteoarthritis: the MOST study. *Annals of the Rheumatic Diseases*. 2011 May;70(5):805–811.

[136] Crema MD, Roemer FW, Marra MD, Burstein D, Gold GE, Eckstein F, et al. Articular cartilage in the knee: current MR imaging techniques and applications in clinical practice and research. *Radiographics*. 2011 Jan;31(1):37–61.

[137] Palmer AJR, Brown CP, McNally EG, Price AJ, Tracey I, Jezzard P, et al. Non-invasive imaging of cartilage in early osteoarthritis. *The Bone & Joint Journal*. 2013 Jun;95-B(6):738–746.

[138] Guermazi A, Niu J, Hayashi D, Roemer FW, Englund M, Neogi T, et al. Prevalence of abnormalities in knees detected by MRI in adults without knee osteoarthritis: population based observational study (Framingham Osteoarthritis Study). *British Medical Journal*. 2012;345:e5339.

[139] Osteoarthritis Initiative;. Available from: `http://www.oai.ucsf.edu/`. [Accessed: April 2014].

[140] Menashe L, Hirko K, Losina E, Kloppenburg M, Zhang W, Li L, et al. The diagnostic performance of MRI in osteoarthritis: A systematic review and meta-analysis. *Osteoarthritis and Cartilage*. 2012;20(1):13–21.

[141] Hunter DJ, Arden N, Conaghan PG, Eckstein F, Gold G, Grainger A, et al. Definition of osteoarthritis on MRI: results of a Delphi exercise. *Osteoarthritis and Cartilage*. 2011 Aug;19(8):963–969.

[142] Guermazi A, Roemer FW, Burstein D, Hayashi D. Why radiography should no longer be considered a surrogate outcome measure for longitudinal

assessment of cartilage in knee osteoarthritis. *Arthritis Research and Therapy*. 2011;13(6):247.

[143] Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Reichmann WM, et al. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. *Osteoarthritis and Cartilage*. 2011 May;19(5):589–605.

[144] Brown CP, Jayadev C, Glyn-Jones S, Carr AJ, Murray DW, Price AJ, et al. Characterization of early stage cartilage degradation using diffuse reflectance near infrared spectroscopy. *Physics in Medicine and Biology*. 2011 Mar;56(7):2299–2307.

[145] Outerbridge RE. The etiology of chondromalacia patellae. *Journal of Bone and Joint Surgery (British)*. 1961 Nov;43-B:752–757.

[146] Brittberg M, Winalski CS. Evaluation of cartilage injuries and repair. *Journal of Bone and Joint Surgery (American)*. 2003;85(SUPPL. 1):58–69.

[147] Brismar BH, Wredmark T, Movin T, Leandersson J, Svensson O. Observer reliability in the arthroscopic classification of osteoarthritis of the knee. *Journal of Bone and Joint Surgery (British)*. 2002 Jan;84(1):42–47.

[148] Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. *Osteoarthritis and Cartilage*. 1997 Mar;5(2):87–97.

[149] Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *The Journal of Rheumatology*. 2000 Jun;27(6):1513–1517.

[150] Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskeletal Disorders*. 2008;9:116.

[151] Cubukcu D, Sarsan A, Alkan H. Relationships between Pain, Function and Radiographic Findings in Osteoarthritis of the Knee: A Cross-Sectional Study. *Arthritis*. 2012;2012:984060.

[152] Wenham CYJ, Conaghan PG. Imaging the painful osteoarthritic knee joint: what have we learned? *Nature clinical practice Rheumatology*. 2009 Mar;5(3):149–158.

[153] White DK, Zhang Y, Niu J, Keysor JJ, Nevitt MC, Lewis CE, et al. Do worsening knee radiographs mean greater chances of severe functional limitation? *Arthritis Care & Research*. 2010 Oct;62(10):1433–1439.

[154] Yusuf E, Kortekaas MC, Watt I, Huizinga TWJ, Kloppenburg M. Do knee abnormalities visualised on MRI explain knee pain in knee osteoarthritis? A systematic review. *Annals of the Rheumatic Diseases*. 2011 Jan;70(1):60–67.

[155] Wluka AE, Wolfe R, Stuckey S, Cicuttini FM. How does tibial cartilage volume relate to symptoms in subjects with knee osteoarthritis? *Annals of the Rheumatic Diseases*. 2004 Mar;63(3):264–268.

[156] Baert IAC, Staes F, Truijen S, Mahmoudian A, Noppe N, Vanderschueren G, et al. Weak associations between structural changes on MRI and symptoms, function and muscle strength in relation to knee osteoarthritis. *Knee surgery, Sports Traumatology, Arthroscopy*. 2013 Feb;.

[157] Driban JB, Sitler MR, Barbe MF, Balasubramanian E. Is osteoarthritis a heterogeneous disease that can be stratified into subsets? *Clinical Rheumatology*. 2010 Feb;29(2):123–131.

[158] Felson DT. Identifying different osteoarthritis phenotypes through epidemiology. *Osteoarthritis and Cartilage*. 2010 May;18(5):601–604.

[159] Bijlsma JWJ, Berenbaum F, Lafeber FPJG. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011 Jun;377(9783):2115–2126.

[160] Conaghan PG. Osteoarthritis in 2012: Parallel evolution of OA phenotypes and therapies. *Nature Reviews Rheumatology*. 2013 Feb;9(2):68–70.

[161] Castañeda S, Roman-Blas JA, Largo R, Herrero-Beaumont G. Osteoarthritis: a progressive disease with changing phenotypes. *Rheumatology*. 2014 Jan;53(1):1–3.

[162] Berenbaum F. Diabetes-induced osteoarthritis: from a new paradigm to a new phenotype. *Annals of the Rheumatic Diseases*. 2011 Jun;70(8):1354–1356.

[163] van der Esch M, Knol DL, Schaffers IC, Reiding DJ, van Schaardenburg D, Knoop J, et al. Osteoarthritis of the knee: multicompartmental or compartmental disease? *Rheumatology*. 2013 Nov;.

[164] Knoop J, van der Leeden M, Thorstensson CA, Roorda LD, Lems WF, Knol DL, et al. Identification of phenotypes with different clinical outcomes in knee osteoarthritis: data from the Osteoarthritis Initiative. *Arthritis Care & Research*. 2011 Nov;63(11):1535–1542.

[165] Felson D, Niu J, Sack B, Aliabadi P, McCullough C, Nevitt MC. Progression of osteoarthritis as a state of inertia. *Annals of the Rheumatic Diseases*. 2012 Jun;.

[166] Gonnella JS, Hornbrook MC, Louis DZ. Staging of disease. A case-mix measurement. *Journal of the American Medical Association*. 1984;251(5):637–644.

[167] Collins DH. *The Pathology of Articular and Spinal Diseases*. Edward Arnold; 1949.

[168] Luyten FP, Denti M, Filardo G, Kon E, Engebretsen L. Definition and classification of early osteoarthritis of the knee. *Knee surgery, Sports Traumatology, Arthroscopy*. 2012 Mar;20(3):401–406.

[169] Kon E, Filardo G, Marcacci M. Early osteoarthritis. *Knee surgery, Sports Traumatology, Arthroscopy*. 2012 Mar;20(3):399–400.

[170] Richette P, Poitou C, Garnero P, Vicaut E, Bouillot JL, Lacorte JM, et al. Benefits of massive weight loss on symptoms, systemic inflammation and cartilage turnover in obese patients with knee osteoarthritis. *Annals of the Rheumatic Diseases*. 2011 Jan;70(1):139–144.

[171] Anandacoomarasamy A, Leibman S, Smith G, Caterson I, Giuffre B, Fransen M, et al. Weight loss in obese people has structure-modifying effects on medial but not on lateral knee articular cartilage. *Annals of the Rheumatic Diseases*. 2012 Jan;71(1):26–32.

[172] Wandel S, Jüni P, Tendal B, Nüesch E, Villiger PM, Welton NJ, et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *British Medical Journal*. 2010;341:c4675.

[173] Wu D, Huang Y, Gu Y, Fan W. Efficacies of different preparations of glucosamine for the treatment of osteoarthritis: a meta-analysis of randomised, double-blind, placebo-controlled trials. *International Journal of Clinical Practice*. 2013 Jun;67(6):585–594.

[174] Reichenbach S, Rutjes AW, Nüesch E, Trelle S, Jüni P. Joint lavage for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews (Online)*. 2010;(5):CD007320.

[175] Katz JN, Brophy RH, Chaisson CE, de Chaves L, Cole BJ, Dahm DL, et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. *New England Journal of Medicine*. 2013 May;368(18):1675–1684.

[176] Virolainen P, Aro HT. High tibial osteotomy for the treatment of osteoarthritis of the knee: a review of the literature and a meta-analysis of follow-up studies. *Archives of Orthopaedic and Trauma Surgery*. 2004 May;124(4):258–261.

[177] Brouwer RW, Raaij van TM, Bierma-Zeinstra SMA, Verhagen AP, Jakma TSC, Verhaar JAN. Osteotomy for treating knee osteoarthritis. *Cochrane Database of Systematic Reviews (Online)*. 2007;(3):CD004019.

[178] Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis and Cartilage*. 2010 Apr;18(4):476–499.

[179] McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI Guidelines for the Non-Surgical Management of Knee Osteoarthritis. *Osteoarthritis and Cartilage*. 2014 Jan;.

[180] National Institute for Health and Care Excellence. Osteoarthritis: The care and management of osteoarthritis in adults; 2008. National Institute for Health and Care Excellence.

[181] American Academy of Orthopaedic Surgeons. Treatment of Osteoarthritis of the Knee: Evidence-Based Guideline. 2nd ed. Evidence-Based Guideline. Rosemont: American Academy of Orthopaedic Surgeons; 2013.

[182] Fernandes L, Hagen KB, Bijlsma JWJ, Andreassen O, Christensen P, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. *Annals of the Rheumatic Diseases*. 2013;72(7):1125–1135.

[183] Moran CG, Horton TC. Total knee replacement: the joint of the decade. A successful operation, for which there's a large unmet need. *British Medical Journal*. 2000 Mar;320(7238):820.

[184] Panel NC. NIH Consensus Statement on total knee replacement December 8-10, 2003. *Journal of Bone and Joint Surgery (American)*. 2004 Jun;86-A(6):1328–1335.

[185] Carr AJ, Robertsson O, Graves S, Price AJ, Arden NK, Judge A, et al. Knee replacement. *Lancet*. 2012 Apr;379(9823):1331–1340.

[186] Hunter DJ. Pharmacologic therapy for osteoarthritis–the era of disease modification. *Nature Reviews Rheumatology*. 2011 Jan;7(1):13–22.

[187] Mobasheri A. The future of osteoarthritis therapeutics: targeted pharmacological therapy. *Current Rheumatology Reports*. 2013 Oct;15(10):364.

[188] Mobasheri A. The future of osteoarthritis therapeutics: emerging biological therapy. *Current Rheumatology Reports*. 2013 Dec;15(12):385.

[189] Qvist P, Bay-Jensen AC, Christiansen C, Dam EB, Pastoureau P, Karsdal MA. The disease modifying osteoarthritis drug (DMOAD): Is it in the horizon? *Pharmacological Research*. 2008 Jul;58(1):1–7.

[190] Felson DT. Osteoarthritis as a disease of mechanics. *Osteoarthritis and Cartilage*. 2013 Jan;21(1):10–15.

[191] Brandt KD. Why should we expect a structure-modifying osteoarthritis drug to relieve osteoarthritis pain? *Annals of the Rheumatic Diseases*. 2011 Jul;70(7):1175–1177.

[192] United States Food & Drug Administration. Clinical Development Programs for Drugs, Devices, and Biological Products Intended for the Treatment of Osteoarthritis (OA); 1999. Available from: `http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatory-%20Information/Guidances/UCM071577.pdf`. [Accessed: April 2014].

[193] European Medicines Agency. Guideline on Clinical Investigation of Medicinal Products Used in the Treatment of Osteoarthritis; 2010. Available from: `http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003440.pdf`. [Accessed: April 2014].

[194] Biomarkers Definitions Working Group . Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*. 2001 Mar;69(3):89–95.

[195] Ferguson MA, Vaidya VS. Biomarkers: An Evolutionary Perspective. In: Vaidya VS, Bonventre JV, editors. *Biomarkers: In Medicine, Drug Discovery, and Environmental Health*. Hoboken: John Wiley & Sons; 2010. p. 1–4.

[196] Lohmander LS, Eyre DR. From biomarker to surrogate outcome to osteoarthritis–what are the challenges? *The Journal of Rheumatology*. 2005 Jun;32(6):1142–1143.

[197] Lohmander SL, Eyre DR. Biochemical Markers as Surrogate End Points of Joint Disease. In: Reid D, Miller C, editors. *Clinical Trials in Rheumatoid Arthritis and Osteoarthritis*. London: Springer-Verlag London Limited; 2008. p. 249–274.

[198] Fox N, Growdon JH. Biomarkers and surrogates. *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics*. 2004 Apr;1(2):181–181.

[199] Kraus VB, Burnett B, Coindreau J, Cottrell S, Eyre D, Gendreau M, et al. Application of biomarkers in the development of drugs intended for the

treatment of osteoarthritis. *Osteoarthritis and Cartilage*. 2011 May;19(5):515–542.

[200] Bauer DC, Hunter DJ, Abramson SB, Attur M, Corr M, Felson D, et al. Classification of osteoarthritis biomarkers: a proposed approach. *Osteoarthritis and Cartilage*. 2006 Aug;14(8):723–727.

[201] Vasan RS. Biomarkers of Cardiovascular Disease: Molecular Basis and Practical Considerations. *Circulation*. 2006 May;113(19):2335–2362.

[202] Dougados M, Hawker G, Lohmander S, Davis AM, Dieppe P, Maillefert JF, et al. OARSI/OMERACT criteria of being considered a candidate for total joint replacement in knee/hip osteoarthritis as an endpoint in clinical trials evaluating potential disease modifying osteoarthritic drugs. *The Journal of Rheumatology*. 2009 Sep;36(9):2097–2099.

[203] Manno RL, Bingham CO, Paternotte S, Gossec L, Halhol H, Giacovelli G, et al. OARSI-OMERACT initiative: defining thresholds for symptomatic severity and structural changes in disease modifying osteoarthritis drug (DMOAD) clinical trials. *Osteoarthritis and Cartilage*. 2012 Feb;20(2):93–101.

[204] Pelletier JP, Cooper C, Peterfy C, Reginster JY, Brandi ML, Bruyère O, et al. What is the predictive value of MRI for the occurrence of knee replacement surgery in knee osteoarthritis? *Annals of the Rheumatic Diseases*. 2013 Oct;72(10):1594–1604.

[205] Simkin PA. Assessing biomarkers in synovial fluid: Consider the kinetics of clearance. *Osteoarthritis and Cartilage*. 2013;21(1):7–9.

[206] Poole R. NIH White Paper: Biomarkers, the Osteoarthritis Initiative. Osteoarthritis Initiative; 2000.

[207] Kraus VB. Do biochemical markers have a role in osteoarhritis diagnosis and treatment? *Best Practice and Research: Clinical Rheumatology*. 2006;20(1):69–80.

[208] Simkin PA, Bassett JE. Cartilage matrix molecules in serum and synovial fluid. *Current Opinion in Rheumatology*. 1995;7(4):346–351.

[209] Atencia LJ, McDevitt CA, Nile WB, Sokoloff L. Cartilage content of an immature dog. *Connective Tissue Research*. 1989;18(4):235–242.

[210] Rayan V, Thonar EJ, Chen LM, Lenz ME, Williams JM. Regional differences in the rise in blood levels of antigenic keratan sulfate and hyaluronan after chymopapain induced knee joint injury. *The Journal of Rheumatology*. 1998 Mar;25(3):521–526.

[211] Xu Q, Sun Xc, Shang Xp, Jiang Hs. Association of CXCL12 levels in synovial fluid with the radiographic severity of knee osteoarthritis. *Journal of Investigative Medicine*. 2012 Aug;60(6):898–901.

[212] Ellabban AS, Kamel SR, Ahmed SS, Osman AM. Receptor activator of nuclear factor kappa B ligand serum and synovial fluid level. A comparative study between rheumatoid arthritis and osteoarthritis. *Rheumatology International*. 2012 Jun;32(6):1589–1596.

[213] Saxne T, Heinegård D. Cartilage oligomeric matrix protein: a novel marker of cartilage turnover detectable in synovial fluid and blood. *British Journal of Rheumatology*. 1992 Sep;31(9):583–591.

[214] Poole AR, Ionescu M, Swan A, Dieppe PA. Changes in cartilage metabolism in arthritis are reflected by altered serum and synovial fluid levels of the cartilage proteoglycan aggrecan: Implications for pathogenesis. *The Journal of Clinical Investigation*. 1994;94(1):25–33.

[215] Misko TP, Radabaugh MR, Highkin M, Abrams M, Friese O, Gallavan R, et al. Characterization of nitrotyrosine as a biomarker for arthritis and joint injury. *Osteoarthritis and Cartilage*. 2013 Jan;21(1):151–156.

[216] Catterall JB, Stabler TV, Flannery CR, Kraus VB. Changes in serum and synovial fluid biomarkers after acute injury (NCT00332254). *Arthritis Research and Therapy*. 2010;12(6):R229.

[217] Giovanni TPS, Golish SR, Palanca A, Hanna LS, Scuderi GJ. Correlation of Intra-Articular Ankle Pathology With Cytokine Biomarkers and Matrix Degradation Products. *Foot & Ankle International*. 2012 Aug;33(08):627–631.

[218] Chockalingam PS, Glasson SS, Lohmander LS. Tenascin-C levels in synovial fluid are elevated after injury to the human and canine joint and correlate with markers of inflammation and matrix degradation. *Osteoarthritis and Cartilage*. 2013 Feb;21(2):339–345.

[219] Kraus VB, Huebner JL, Fink C, King JB, Brown S, Vail TP, et al. Urea as a passive transport marker for arthritis biomarker studies. *Arthritis and Rheumatism*. 2002;46(2):420–427.

[220] Kraus V, Stabler T, Kong S, Varju G, McDaniel G. Measurement of synovial fluid volume using urea. *Osteoarthritis and Cartilage*. 2007 Jan;15(10):1217–1220.

[221] Dahlberg L, Friden T, Roos H, Lark MW, Lohmander LS. A longitudinal study of cartilage matrix metabolism in patients with cruciate ligament rupture - Synovial fluid concentrations of aggrecan fragments, stromelysin-1 and tissue inhibitor of metalloproteinase-1. *British Journal of Rheumatology*. 1994;33(12):1107–1111.

[222] Lohmander LS, Dahlberg L, Eyre D, Lark M, Thonar EJ, Ryd L. Longitudinal and cross-sectional variability in markers of joint metabolism in patients with knee pain and articular cartilage abnormalities. *Osteoarthritis and Cartilage*. 1998 Sep;6(5):351–361.

[223] van Spil WE, Degroot J, Lems WF, Oostveen JCM, Lafeber FPJG. Serum and urinary biochemical markers for knee and hip-osteoarthritis: a systematic review applying the consensus BIPED criteria. *Osteoarthritis and Cartilage*. 2010 May;18(5):605–612.

[224] Qvist P, Christiansen C, Karsdal MA, Madsen SH, Sondergaard BC, Bay-Jensen AC. Application of biochemical markers in development of drugs for treatment of osteoarthritis. *Biomarkers*. 2010 Feb;15(1):1–19.

[225] Patra D, Sandell LJ. Recent advances in biomarkers in osteoarthritis. *Current Opinion in Rheumatology*. 2011 Sep;23(5):465–470.

[226] Mobasheri A, Henrotin Y. Biomarkers of osteoarthritis: a review of recent research progress on soluble biochemical markers, published patents and areas for future development. *Recent Patents on Biomarkers*. 2011;.

[227] De Ceuninck F, Sabatini M, Pastoureau P. Recent progress toward biomarker identification in osteoarthritis. *Drug Discovery Today*. 2011 May;16(9-10):443–449.

[228] Rousseau JC, Garnero P. Biological markers in osteoarthritis. *Bone*. 2012 Aug;51(2):265–277.

[229] Mobasheri A. Osteoarthritis year 2012 in review: biomarkers. *Osteoarthritis and Cartilage*. 2012 Jul;.

[230] Attur M, Krasnokutsky-Samuels S, Samuels J, Abramson SB. Prognostic biomarkers in osteoarthritis. *Current Opinion in Rheumatology*. 2013 Jan;25(1):136–144.

[231] Lafeber FPJG, van Spil WE. Osteoarthritis year 2013 in review: biomarkers; reflecting before moving forward, one step at a time. *Osteoarthritis and Cartilage*. 2013 Oct;21(10):1452–1464.

[232] Lotz M, Martel-Pelletier J, Christiansen C, Brandi ML, Bruyère O, Chapurlat R, et al. Value of biomarkers in osteoarthritis: current status and perspectives. *Annals of the Rheumatic Diseases*. 2013 Aug;.

[233] Garnero P, Rousseau JC, Delmas PD. Molecular basis and clinical use of biochemical markers of bone, cartilage, and synovium in joint diseases. *Arthritis and Rheumatism*. 2000 May;43(5):953–968.

[234] Ruiz-Romero C, Blanco F. Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis and Cartilage*. 2010 Jan;18(4):500–509.

[235] Adams SB, Setton LA, Nettles DL. The role of metabolomics in osteoarthritis research. *Journal of the American Academy of Orthopaedic Surgeons*. 2013 Jan;21(1):63–64.

[236] Felson DT, Lohmander LS. Whither osteoarthritis biomarkers? *Osteoarthritis and Cartilage*. 2009 Apr;17(4):419–422.

[237] Jenkins M, Flynn A, Smart T, Harbron C, Sabin T, Ratnayake J, et al. A statistician's perspective on biomarkers in drug development. *Pharmaceutical Statistics*. 2011 Dec;10(6):494–507.

[238] Williams FMK. Biomarkers: in combination they may do better. *Arthritis Research and Therapy*. 2009;11(5):130.

[239] Meulenbelt I, Kloppenburg M, Kroon HM, Houwing-Duistermaat JJ, Garnero P, Hellio-Le Graverand MP, et al. Clusters of biochemical markers are associated with radiographic subtypes of osteoarthritis (OA) in subject with familial OA at multiple sites. The GARP study. *Osteoarthritis and Cartilage*. 2007 Apr;15(4):379–385.

[240] van Spil WE, Jansen NWD, Bijlsma JWJ, Reijman M, Degroot J, Welsing PMJ, et al. Clusters within a wide spectrum of biochemical markers for osteoarthritis: data from CHECK, a large cohort of individuals with very early symptomatic osteoarthritis. *Osteoarthritis and Cartilage*. 2012 Jul;20(7):745–754.

[241] Beekhuizen M, Gierman LM, van Spil WE, van Osch GJVM, Huizinga TWJ, Saris DBF, et al. An explorative study comparing levels of soluble mediators in control and osteoarthritic synovial fluid. *Osteoarthritis and Cartilage*. 2013 Jul;21(7):918–922.

[242] Cibere J, Zhang H, Garnero P, Poole AR, Lobanok T, Saxne T, et al. Association of biomarkers with pre-radiographically defined and radiographically defined knee osteoarthritis in a population-based study. *Arthritis and Rheumatism*. 2009;60(5):1372–1380.

[243] Hirata S, Dirven L, Shen Y, Centola M, Cavet G, Lems WF, et al. A multi-biomarker score measures rheumatoid arthritis disease activity in the BeSt study. *Rheumatology*. 2013 Jul;52(7):1202–1207.

[244] Centola M, Cavet G, Shen Y, Ramanujan S, Knowlton N, Swan KA, et al. Development of a multi-biomarker disease activity test for rheumatoid arthritis. *PloS One*. 2013;8(4):e60635.

[245] Jiang M, Chen T, Feng H, Zhang Y, Li L, Zhao A, et al. Serum metabolic signatures of four types of human arthritis. *Journal of Proteome Research*. 2013;12(8):3769–3779.

[246] Han MY, Dai JJ, Zhang Y, Lin Q, Jiang M, Xu XY, et al. Identification of osteoarthritis biomarkers by proteomic analysis of synovial fluid. *Journal of International Medical Research*. 2012;40(6):2243–2250.

[247] Dam EB, Loog M, Christiansen C, Byrjalsen I, Folkesson J, Nielsen M, et al. Identification of progressors in osteoarthritis by combining biochemical and MRI-based markers. *Arthritis Research and Therapy*. 2009;11(4):R115.

[248] FNIH Osteoarthritis Biomarkers Consortium Project;. Available from: `http://oarsi.org/sites/default/files/fnih_oa_biomarkers_project_update_0.pdf`. [Accessed: April 2014].

[249] STARD Initiative;. Available from: `http://www.stard-statement.org`. [Accessed: April 2014].

[250] White SH, Ludkowski PF, Goodfellow JW. Anteromedial osteoarthritis of the knee. *Journal of Bone and Joint Surgery (British)*. 1991 Jul;73(4):582–586.

[251] Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, et al. Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. *Quality of Life Research*. 2013 Mar;.

[252] Riddle DL, Jiranek WA, Hull JR. Validity and reliability of radiographic knee osteoarthritis measures by arthroplasty surgeons. *Orthopedics*. 2013 Jan;36(1):e25–32.

[253] Osteoarthritis Initiative. Project 06 and Project 14: Test-Retest Reliability Analysis of Semi-quantitative Readings of Knee Radiographs. Summary Report. Osteoarthritis Initiative; 2011.

[254] Osteoarthritis Initiative. Project 15: Test-Retest Reliability of Semi-quantitative Readings from Knee Radiographs. Osteoarthritis Initiative; 2012.

[255] Wild D, editor. The Immunoassay Handbook. 4th ed. Theory and Applications of Ligand Binding, ELISA and Related Techniques. Oxford: Elsevier; 2013.

[256] Luminex xMAP Technology;. Available from: `http://www.luminexcorp.com/TechnologiesScience/xMAPTechnology/`. [Accessed: April 2014].

[257] Meso Scale Discovery;. Available from: `http://www.mesoscale.com/CatalogSystemWeb/WebRoot/technology.htm`. [Accessed: April 2014].

[258] Gottschalk PG, Dunn JR. The five-parameter logistic: a characterization and comparison with the four-parameter logistic. *Analytical Biochemistry*. 2005 Aug;343(1):54–65.

[259] Liu A. 10 Tips for Luminex/Bio-Plex Data Analysis; 2009. Available from: `http://www.miraibio.com/blog/2009/02/tips-for-luminex-data-analysis/`. [Accessed: April 2014].

[260] DeSilva B, Smith W, Weiner R, Kelley M, Smolec J, Lee B, et al. Recommendations for the Bioanalytical Method Validation of Ligand-binding Assays to Support Pharmacokinetic Assessments of Macromolecules. *Pharmaceutical Research*. 2003 Nov;20(11):1885–1900.

[261] Smolec J, DeSilva B, Smith W, Weiner R, Kelly M, Lee B, et al. Bioanalytical method validation for macromolecules in support of pharmacokinetic studies. *Pharmaceutical Research*. 2005 Sep;22(9):1425–1431.

[262] Viswanathan CT, Bansal S, Booth B, DeStefano AJ, Rose MJ, Sailstad J, et al.; Food; Center for Drug Evaluation Drug Administration; Maryland 20857 USA. Research. Quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. *Pharmaceutical Research*. 2007 Oct;24(10):1962–1973.

[263] Valentin MA, Ma S, Zhao A, Legay F, Avrameas A. Validation of immunoassay for protein biomarkers: Bioanalytical study plan implementation to support pre-clinical and clinical studies. *Journal of Pharmaceutical and Biomedical Analysis*. 2011 Jul;55(5):869–877.

[264] HOLM S. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*. 1979;6(2):65–70.

[265] Johnston JP. The viscosity of normal and pathological human synovial fluids. *The Biochemical Journal*. 1955 Apr;59(4):633–637.

[266] Reimann I. Pathological human synovial fluids. Viscosity and boundary lubricating properties. *Clinical Orthopaedics and Related Research*. 1976 Sep;No.119(119):237–241.

[267] Song SY, Han YD, Hong SY, Kim K, Yang SS, Min BH, et al. Chip-based cartilage oligomeric matrix protein detection in serum and synovial fluid for osteoarthritis diagnosis. *Analytical Biochemistry*. 2012 Jan;420(2):139–146.

[268] de Jager W, Hoppenreijs EPAH, Wulffraat NM, Wedderburn LR, Kuis W, Prakken BJ. Blood and synovial fluid cytokine signatures in patients with juvenile idiopathic arthritis: a cross-sectional study. *Annals of the Rheumatic Diseases*. 2007 May;66(5):589–598.

[269] Mateos J, Lourido L, Fernández-Puente P, Calamia V, Fernández-López C, Oreiro N, et al. Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC–MALDI TOF/TOF. *Journal of Proteomics*. 2012 Jan;p. 1–10.

[270] Sugiuchi H, Ando Y, Manabe M, Nakamura E, Mizuta H, Nagata S, et al. Measurement of total and differential white blood cell counts in synovial fluid by means of an automated hematology analyzer. *Journal of Laboratory and Clinical Medicine*. 2005 Jul;146(1):36–42.

[271] Ekmann A, Rigdal ML, Gröndahl G. Automated counting of nucleated cells in equine synovial fluid without and with hyaluronidase pretreatment. *Veterinary Clinical Pathology*. 2010 Mar;39(1):83–89.

[272] Platt D, Pigman W, Holley HL. The effect of testicular hyaluronidase on the hyaluronic acid complexes of human synovial fluid. *The Biochemical Journal*. 1961 Jun;79:605–608.

[273] Rinaudo M, Rozand Y, Mathieu P, Conrozier T. Role of Different Pre-Treatments on Composition and Rheology of Synovial Fluids. *Polymers*. 2009 Dec;1(1):16–34.

[274] Banks RE. Measurement of cytokines in clinical samples using immunoassays: problems and pitfalls. *Critical Reviews in Clinical Laboratory Sciences*. 2000 Apr;37(2):131–182.

[275] Efthimiadis A, Spanevello A, Hamid Q, Kelly MM, Linden M, Louis R, et al. Methods of sputum processing for cell counts, immunocytochemistry and in situ hybridisation. *The European Respiratory Journal*. 2002 Sep;20(Suppl 37):19s–23s.

[276] Kelly MM, Keatings V, Leigh R, Peterson C, Shute J, Venge P, et al. Analysis of fluid phase mediators. *The European Respiratory Journal*. 2002 Jul;20(Suppl. 37):24S–39s.

[277] Malyak M, Joslin FG, Verderber EL, Eisenberg SP, Arend WP. IL-1ra ELISA: reduction and alkylation of synovial fluid eliminates interference by IgM rheumatoid factors. *Journal of Immunological Methods*. 1991 Jul;140(2):281–288.

[278] Smith MA, Bains SK, Betts JC, Choy EH, Zanders ED. Use of two-dimensional gel electrophoresis to measure changes in synovial fluid proteins from patients with rheumatoid arthritis treated with antibody to CD4. *Clinical and Diagnostic Laboratory Immunology*. 2001 Jan;8(1):105–111.

[279] Yamagiwa H, Sarkar G, Charlesworth MC, McCormick DJ, Bolander ME. Two-dimensional gel electrophoresis of synovial fluid: method for detecting candidate protein markers for osteoarthritis. *Journal of Orthopaedic Science*. 2003;8(4):482–490.

[280] Woolhouse IS, Bayley DL, Stockley RA. Effect of sputum processing with dithiothreitol on the detection of inflammatory mediators in chronic bronchitis and bronchiectasis. *Thorax*. 2002 Aug;57(8):667–671.

[281] Jacobson JW, Oliver KG, Weiss C, Kettman J. Analysis of individual data from bead-based assays ("bead arrays"). *Cytometry Part A*. 2006 May;69(5):384–390.

[282] Cook D, Giron J, Sekella P. Evaluation of Luminex immunoassay precision: comparison of MFI precision when collecting 50 versus 100 bead events per region. Invitrogen Immunology Systems; 2008.

[283] Moreno MJ, Clayburne G, Schumacher HR. Processing of noninflammatory synovial fluids with hyaluronidase for cytospin preparations improves the accuracy of differential counts. *Diagnostic Cytopathology*. 2000 Apr;22(4):256–258.

[284] Levinson SS, Miller JJ. Towards a better understanding of heterophile (and the like) antibody interference with modern immunoassays. *Clinica Chimica Acta*. 2002 Nov;325(1-2):1–15.

[285] de Jager W, Rijkers GT. Solid-phase and bead-based cytokine immunoassay: a comparison. *Methods*. 2006 Apr;38(4):294–303.

[286] Bjerner J, Børmer OP, Nustad K. The war on heterophilic antibody interference. *Clinical Chemistry*. 2005 Jan;51(1):9–11.

[287] de Jager W, Prakken BJ, Bijlsma JWJ, Kuis W, Rijkers GT. Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. *Journal of Immunological Methods*. 2005 May;300(1-2):124–135.

[288] DeForge LE, Loyet KM, Delarosa D, Chinn J, Zamanian F, Chuntharapai A, et al. Evaluation of heterophilic antibody blocking agents in reducing false positive interference in immunoassays for IL-17AA, IL-17FF, and IL-17AF. *Journal of Immunological Methods*. 2010 Oct;362(1-2):70–81.

[289] Chowdhury F, Williams A, Johnson P. Validation and comparison of two multiplex technologies, Luminex and Mesoscale Discovery, for human cytokine profiling. *Journal of Immunological Methods*. 2009 Jan;340(1):55–64.

[290] Sourial S, Marcusson-Ståhl M, Cederbrant K. Meso scale discovery and Luminex comparative analysis of calbindin D28K. *Journal of Biomedicine and Biotechnology*. 2009;2009.

[291] Fu Q, Zhu J, Van Eyk JE. Comparison of multiplex immunoassay platforms. *Clinical Chemistry*. 2010 Feb;56(2):314–318.

[292] Dabitao D, Margolick JB, Lopez J, Bream JH. Multiplex measurement of proinflammatory cytokines in human serum: comparison of the Meso Scale Discovery electrochemiluminescence assay and the Cytometric Bead Array. *Journal of Immunological Methods*. 2011 Sep;372(1-2):71–77.

[293] Breen EC, Reynolds SM, Cox C, Jacobson LP, Magpantay L, Mulder CB, et al. Multisite comparison of high-sensitivity multiplex cytokine assays. *Clinical and Vaccine Immunology*. 2011 Aug;18(8):1229–1242.

[294] Thompson DK, Huffman KM, Kraus WE, Kraus VB. Critical appraisal of four IL-6 immunoassays. *PloS One*. 2012;7(2):e30659.

[295] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*. 1979 Mar;86(2):420–428.

[296] Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989 Mar;45(1):255–268.

[297] Lin LIK. Erratum: A note on the concordance correlation coefficient. *Biometrics*. 2000;56(1):324–325.

[298] Linnet K. Evaluation of regression procedures for methods comparison studies. *Clinical Chemistry*. 1993 Mar;39(3):424–432.

[299] Martin RF. General deming regression for estimating systematic bias and its confidence interval in method-comparison studies. *Clinical Chemistry*. 2000 Jan;46(1):100–104.

[300] Ludbrook J. Linear regression analysis for comparing two measurers or methods of measurement: But which regression? *Clinical and Experimental Pharmacology and Physiology*. 2010;37(7):692–699.

[301] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*. 1999 Jun;8(2):135–160.

[302] Ludbrook J. Confidence in Altman-Bland plots: A critical review of the method of differences. *Clinical and Experimental Pharmacology and Physiology*. 2010;37(2):143–149.

[303] McGraw KO, Wong SP. Forming Inferences about Some Intraclass Correlation Coefficients. *Psychological Methods*. 1996;1(1):30–46.

[304] McBride GB. A proposal for strength-of agreement criteria for Lin's Concordance Correlation Coefficient. NIWA Client Report: HAM2005-062; 2005.

[305] de Jager W, te Velthuis H, Prakken BJ, Kuis W, Rijkers GT. Simultaneous detection of 15 human cytokines in a single sample of stimulated peripheral blood mononuclear cells. *Clinical and Diagnostic Laboratory Immunology*. 2003 Jan;10(1):133–139.

[306] duPont NC, Wang K, Wadhwa PD, Culhane JF, Nelson EL. Validation and comparison of luminex multiplex cytokine analysis kits with ELISA: Determinations of a panel of nine cytokines in clinical sample culture supernatants. *Journal of Reproductive Immunology*. 2005 Aug;66(2):175–191.

[307] Fichorova RN, Richardson-Harman N, Alfano M, Belec L, Carbonneil C, Chen S, et al. Biological and technical variables affecting immunoassay recovery of cytokines from human serum and simulated vaginal fluid: a multicenter study. *Analytical Chemistry*. 2008 Jun;80(12):4741–4751.

[308] Kang JH, Vanderstichele H, Trojanowski JQ, Shaw LM. Simultaneous analysis of cerebrospinal fluid biomarkers using microsphere-based xMAP multiplex technology for early detection of Alzheimer's disease. *Methods*. 2012 Apr;56(4):484–493.

[309] Leng SX, McElhaney JE, Walston JD, Xie D, Fedarko NS, Kuchel GA. ELISA and multiplex technologies for cytokine measurement in inflammation and aging research. *Journals of Gerontology Series A, Biological Sciences and Medical sciences*. 2008 Aug;63(8):879–884.

[310] Fan A, Cao Z, Li H, Kai M, Lu J. Chemiluminescence platforms in immunoassay and DNA analyses. *Analytical Sciences*. 2009 May;25(5):587–597.

[311] Rudy JL. Differentiating between sensitivity and limit of detection. *Clinical Chemistry*. 1989;35(3):509.

[312] Prabhakar U, Eirikis E, Reddy M, Silvestro E, Spitz S, Pendley C, et al. Validation and comparative analysis of a multiplexed assay for the simultaneous quantitative measurement of Th1/Th2 cytokines in human serum and human peripheral blood mononuclear cell culture supernatants. *Journal of Immunological Methods*. 2004 Aug;291(1-2):27–38.

[313] Guidance for Industry: Bioanalytical Method Validation . US Department of Health; 2001.

[314] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310.

[315] Ledur A, Fitting C, David B, Hamberger C, Cavaillon JM. Variable estimates of cytokine levels produced by commercial ELISA kits: results using international cytokine standards. *Journal of Immunological Methods*. 1995 Oct;186(2):171–179.

[316] Aziz N, Nishanian P, Mitsuyasu R, Detels R, Fahey JL. Variables that affect assays for plasma cytokines and soluble activation markers. *Clinical and Diagnostic Laboratory Immunology*. 1999 Jan;6(1):89–95.

[317] Roux-Lombard P, Steiner G. Preliminary report on cytokine determination in human synovial fluids: a consensus study of the European Workshop for Rheumatology Research. The Cytokine Consensus Study Group of the European Workshop for Rheumatology Research. *Clinical and Experimental Rheumatology*. 1992 Sep;10(5):515–520.

[318] Tammemagi MC, Frank JW, LeBlanc M, Artsob H, Streiner DL. Methodological Issues in Assessing Reproducibility - a Comparative-Study of Various Indexes of Reproducibility Applied to Repeat Elisa Serologic Tests for Lyme-Disease. *Journal of Clinical Epidemiology*. 1995;48(9):1123–1132.

[319] Chen CC, Barnhart HX. Comparison of ICC and CCC for assessing agreement for data without and with replications. *Computational Statistics and Data Analysis*. 2008;53(2):554–564.

[320] Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Computers in Biology and Medicine*. 1990;20(5):337–340.

[321] Prabhakar U, Eirikis E, Davis HM. Simultaneous quantification of proinflammatory cytokines in human plasma using the LabMAP assay. *Journal of Immunological Methods*. 2002 Feb;260(1-2):207–218.

[322] Elshal MF, McCoy JP. Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA. *Methods*. 2006 Apr;38(4):317–323.

[323] Schipper HS, de Jager W, van Dijk MEA, Meerding J, Zelissen PMJ, Adan RA, et al. A multiplex immunoassay for human adipokine profiling. *Clinical Chemistry*. 2010 Aug;56(8):1320–1328.

[324] Schipke CG, Prokop S, Heppner FL, Heuser I, Peters O. Comparison of immunosorbent assays for the quantification of biomarkers for Alzheimer's disease in human cerebrospinal fluid. *Dementia and Geriatric Cognitive Disorders*. 2011;31(2):139–145.

[325] Goldblatt D, Ashton L, Zhang Y, Antonello J, Marchese RD. Comparison of a new multiplex binding assay versus the enzyme-linked immunosorbent assay for measurement of serotype-specific pneumococcal capsular polysaccharide IgG. *Clinical and Vaccine Immunology*. 2011 Oct;18(10):1744–1751.

[326] Regeniter A, Kuhle J, Baumann T, Sollberger M, Herdener M, Kunze U, et al. Biomarkers of dementia: comparison of electrochemiluminescence results and reference ranges with conventional ELISA. *Methods*. 2012 Apr;56(4):494–499.

[327] de Koning L, Liptak C, Shkreta A, Bradwin G, Hu FB, Pradhan AD, et al. A multiplex immunoassay gives different results than singleplex immunoassays which may bias epidemiologic associations. *Clinical Biochemistry*. 2012 Jul;45(10-11):848–851.

[328] Hildesheim A, Ryan RL, Rinehart E, Nayak S, Wallace D, Castle PE, et al. Simultaneous measurement of several cytokines using small volumes of biospecimens. *Cancer Epidemiology, Biomarkers & Prevention*. 2002 Nov;11(11):1477–1484.

[329] Ray CA, Bowsher RR, Smith WC, Devanarayan V, Willey MB, Brandt JT, et al. Development, validation, and implementation of a multiplex immunoassay for the simultaneous determination of five cytokines in human serum. *Journal of Pharmaceutical and Biomedical Analysis*. 2005 Jan;36(5):1037–1044.

[330] Liu MY, Xydakis AM, Hoogeveen RC, Jones PH, Smith EO, Nelson KW, et al. Multiplexed analysis of biomarkers related to obesity and the metabolic syndrome in human plasma, using the Luminex-100 system. *Clinical Chemistry*. 2005 Jul;51(7):1102–1109.

[331] Khan SS, Smith MS, Reda D, Suffredini AF, McCoy JP. Multiplex bead array assays for detection of soluble cytokines: comparisons of sensitivity and quantitative values among kits from multiple manufacturers. *Cytometry Part B*. 2004 Sep;61(1):35–39.

[332] Pickering JW, Martins TB, Schroder MC, Hill HR. Comparison of a multiplex flow cytometric assay with enzyme-linked immunosorbent assay for auantitation of antibodies to tetanus, diphtheria, and Haemophilus influenzae Type b. *Clinical and Diagnostic Laboratory Immunology*. 2002 Jul;9(4):872–876.

[333] Toedter G, Hayden K, Wagner C, Brodmerkel C. Simultaneous detection of eight analytes in human serum by two commercially available platforms for multiplex cytokine analysis. *Clinical and Vaccine Immunology*. 2008 Jan;15(1):42–48.

[334] Skogstrand K, Thorsen P, Nørgaard-Pedersen B, Schendel DE, Sørensen LC, Hougaard DM. Simultaneous measurement of 25 inflammatory markers and neurotrophins in neonatal dried blood spots by immunoassay with xMAP technology. *Clinical Chemistry*. 2005 Oct;51(10):1854–1866.

[335] Kofoed K, Schneider UV, Scheel T, Andersen O, Eugen-Olsen J. Development and validation of a multiplex add-on assay for sepsis biomarkers using xMAP technology. *Clinical Chemistry*. 2006 Jul;52(7):1284–1293.

[336] Djoba Siawaya JF, Roberts T, Babb C, Black G, Golakai HJ, Stanley K, et al. An evaluation of commercial fluorescent bead-based luminex cytokine assays. *PloS One*. 2008;3(7):e2535.

[337] Thrailkill KM, Moreau CS, Cockrell G, Simpson P, Goel R, North P, et al. Physiological matrix metalloproteinase concentrations in serum during childhood and adolescence, using Luminex Multiplex technology. *Clinical Chemistry and Laboratory Medicine*. 2005;43(12):1392–1399.

[338] Dossus L, Becker S, Achaintre D, Kaaks R, Rinaldi S. Validity of multiplex-based assays for cytokine measurements in serum and plasma from "non-diseased" subjects: Comparison with ELISA. *Journal of Immunological Methods*. 2009 Oct;350(1-2):125–132.

[339] Richens JL, Urbanowicz RA, Metcalf R, Corne J, O'Shea P, Fairclough L. Quantitative validation and comparison of multiplex cytokine kits. *Journal of Biomolecular Screening*. 2010 Jun;15(5):562–568.

[340] Ellington AA, Kullo IJ, Bailey KR, Klee GG. Antibody-Based Protein Multiplex Platforms: Technical and Operational Challenges. *Clinical Chemistry*. 2010 Jan;56(2):186–193.

[341] Kingsmore SF. Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nature reviews Drug discovery*. 2006 Apr;5(4):310–320.

[342] de Jager W, Bourcier K, Rijkers GT, Prakken BJ, Seyfert-Margolis V. Prerequisites for cytokine measurements in clinical trials with multiplex immunoassays. *BMC Immunology*. 2009;10:52.

[343] Porter AE, Auth J, Prince M, Ghidini A, Brenneman DE, Spong CY. Optimization of cytokine stability in stored amniotic fluid. *American Journal of Obstetrics and Gynecology*. 2001 Aug;185(2):459–462.

[344] Cruz Ruiz M, Recio Quijano F, López Cortés LF, Hebles Duvison M, Vázquez Rubio R. Determination of shelf life and activation energy for tumor necrosis factor-

alpha in cerebrospinal fluid samples. *Clinical Chemistry*. 1996 May;42(5):670–674.

[345] Bielsa S, Davies HE, Davies RJO, Taylor A, Lee YCG. Reproducibility and reliability of pleural fluid cytokine measurements. *The European Respiratory Journal*. 2009 Oct;34(4):1001–1003.

[346] Hosnijeh FS, Krop EJM, Portengen L, Rabkin CS, Linseisen J, Vineis P, et al. Stability and reproducibility of simultaneously detected plasma and serum cytokine levels in asymptomatic subjects. *Biomarkers*. 2010 Mar;15(2):140–148.

[347] Thavasu PW, Longhurst S, Joel SP, Slevin ML, Balkwill FR. Measuring cytokine levels in blood. Importance of anticoagulants, processing, and storage conditions. *Journal of Immunological Methods*. 1992 Aug;153(1-2):115–124.

[348] Kenis G, Teunissen C, De Jongh R, Bosmans E, Steinbusch H, Maes M. Stability of interleukin 6, soluble interleukin 6 receptor, interleukin 10 and CC16 in human serum. *Cytokine*. 2002 Sep;19(5):228–235.

[349] Flower L, Ahuja RH, Humphries SE, Mohamed-Ali V. Effects of sample handling on the stability of interleukin 6, tumour necrosis factor-alpha and leptin. *Cytokine*. 2000 Nov;12(11):1712–1716.

[350] Bottomley N, Kendrick B, Ferguson J, Al-Ali S, Gill H, Beard D, et al. What proportion of knee surgery practice is osteoarthritis and are we treating it correctlyy? *Journal of Bone and Joint Surgery (British)*. 2012;94-B(SUPP IX):16.

[351] Last RJ. Specimens from the Hunterian collection. 4. The synovial cavity of the knee joint (Specimen S 110 A). *Journal of Bone and Joint Surgery (British)*. 1951 Aug;33-B(3):442–445.

[352] Knight AD, Levick JR. Physiological Compartmentation of Fluid within the Synovial Cavity of the Rabbit Knee. *Journal of Physiology*. 1982 Jan;331:1–15.

[353] Wood L, Ferrell WR. Fluid compartmentation and articular mechanoreceptor discharge in the cat knee joint. *Quarterly journal of experimental physiology (Cambridge, England)*. 1985 Jul;70(3):329–335.

[354] Pedowitz RA, Gershuni DH, Crenshaw AG, Petras SL, Danzig LA, Hargens AR. Intraarticular pressure during continuous passive motion of the human knee. *Journal of Orthopaedic Research*. 1989;7(4):530–537.

[355] McCarty WJ, Masuda K, Sah RL. Fluid movement and joint capsule strains due to flexion in rabbit knees. *Journal of Biomechanics*. 2011 Nov;44(16):2761–2767.

[356] Kerkhof HJM, Meulenbelt I, Akune T, Arden NK, Aromaa A, Bierma-Zeinstra SMA, et al. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. *Osteoarthritis and Cartilage*. 2011 Mar;19(3):254–264.

[357] Roemer FW, Guermazi A, Niu J, Zhang Y, Mohr A, Felson DT. Prevalence of magnetic resonance imaging-defined atrophic and hypertrophic phenotypes of knee osteoarthritis in a population-based cohort. *Arthritis and Rheumatism*. 2012;64(2):429–437.

[358] Ledingham J, Regan M, Jones A, Doherty M. Radiographic patterns and associations of osteoarthritis of the knee in patients referred to hospital. *Annals of the Rheumatic Diseases*. 1993 Jul;52(7):520–526.

[359] van Osch GJVM, van der Kraan PM, Van Valburg AA, van den Berg WB. The relation between cartilage damage and osteophyte size in a murine model for osteoarthritis in the knee. *Rheumatology International*. 1996;16(3):115–119.

[360] Young AA, Smith MM, Smith SM, Cake MA, Ghosh P, Read RA, et al. Regional assessment of articular cartilage gene expression and small proteoglycan metabolism in an animal model of osteoarthritis. *Arthritis Research and Therapy*. 2005;7(4):R852–61.

[361] Grainger AJ, Rhodes LA, Keenan AM, Emery P, Conaghan PG. Quantifying peri-meniscal synovitis and its relationship to meniscal pathology in osteoarthritis of the knee. *European Radiology*. 2007;17(1):119–124.

[362] Roemer FW, Kassim Javaid M, Guermazi A, Thomas M, Kiran A, Keen R, et al. Anatomical distribution of synovitis in knee osteoarthritis and its association

with joint effusion assessed on non-enhanced and contrast-enhanced MRI. *Osteoarthritis and Cartilage*. 2010;18(10):1269–1274.

[363] Saito T, Koshino T. Distribution of neuropeptides in synovium of the knee with osteoarthritis. *Clinical Orthopaedics and Related Research*. 2000;(376):172–182.

[364] Sutton S, Clutterbuck A, Harris P, Gent T, Freeman S, Foster N, et al. The contribution of the synovium, synovial derived inflammatory cytokines and neuropeptides to the pathogenesis of osteoarthritis. *Veterinary Journal*. 2009 Jan;179(1):10–24.

[365] Nord RM, Badowski N, Elkins C, Alley M, Harris AHS, Dragoo JL. Optimizing lavage during knee arthroscopy: A three-dimensional MRI study. *Journal of Magnetic Resonance Imaging*. 2013 Jan;37(1):201–207.

[366] Clough SB, Read HE, Metzner AB, Behn VC. Diffusion in slurries and in non-Newtonian fluids. *AIChE Journal*. 1962 Jun;8(3):346–350.

[367] Hadler NM. Synovial fluids facilitate small solute diffusivity. *Annals of the Rheumatic Diseases*. 1980;39(6):580–585.

[368] Hron J, Málek J, Pustějovská P, Rajagopal KR. On the Modeling of the Synovial Fluid. *Advances in Tribology*. 2010;2010:1–12.

[369] Chhabra RP. Non-Newtonian Fluids: An Introduction. In: Deshpande A, Krishnan JM, Kumar PBS, editors. *Rheology of Complex Fluids*. Springer; 2010. p. 1–33.

[370] Roos H, Dahlberg L, Hoerrner LA, Lark MW, Thonar EJ, Shinmei M, et al. Markers of cartilage matrix metabolism in human joint fluid and serum: the effect of exercise. *Osteoarthritis and Cartilage*. 1995 Mar;3(1):7–14.

[371] Criscione LG, Elliott AL, Stabler T, Jordan JM, Pieper CF, Kraus VB. Variation of serum hyaluronan with activity in individuals with knee osteoarthritis. *Osteoarthritis and Cartilage*. 2005;13(9):837–840.

[372] Mündermann A, Dyrby CO, Andriacchi TP, King KB. Serum concentration of cartilage oligomeric matrix protein (COMP) is sensitive to physiological cyclic loading in healthy adults. *Osteoarthritis and Cartilage*. 2005 Jan;13(1):34–38.

[373] Andersson MLE, Petersson IF, Karlsson KE, Jonsson EN, Månsson B, Heinegård D, et al. Diurnal variation in serum levels of cartilage oligomeric matrix protein in patients with knee osteoarthritis or rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2006 Nov;65(11):1490–1494.

[374] Kong SY, Stabler TV, Criscione LG, Elliott AL, Jordan JM, Kraus VB. Diurnal variation of serum and urine biomarkers in patients with radiographic knee osteoarthritis. *Arthritis and Rheumatism*. 2006 Aug;54(8):2496–2504.

[375] Gordon CD, Stabler TV, Kraus VB. Variation in osteoarthritis biomarkers from activity not food consumption. *Clinica Chimica Acta*. 2008;398(1-2):21–26.

[376] Quintana DJ, Garnero P, Huebner JL, Charni-Ben Tabassi N, Kraus VB. PIIANP and HELIXII diurnal variation. *Osteoarthritis and Cartilage*. 2008;16(10):1192–1195.

[377] Mündermann A, King KB, Smith RL, Andriacchi TP. Change in serum COMP concentration due to ambulatory load is not related to knee OA status. *Journal of Orthopaedic Research*. 2009 Nov;27(11):1408–1413.

[378] Lohmander LS, Atley LM, Pietka TA, Eyre DR. The release of crosslinked peptides from type II collagen into human synovial fluid is increased soon after joint injury and in osteoarthritis. *Arthritis and Rheumatism*. 2003 Nov;48(11):3130–3139.

[379] Nakajima A, Nakagawa K, Aoki Y, Sonobe M, Shibata Y, Yamazaki M, et al. Changes in synovial fluid biochemical markers following arthroscopic surgery in patients with knee osteoarthritis. *Rheumatology International*. 2012;p. 1–6.

[380] Blennow K, Zetterberg H, Minthon L, Lannfelt L, Strid S, Annas P, et al. Longitudinal stability of CSF biomarkers in Alzheimer's disease. *Neuroscience Letters*. 2007 May;419(1):18–22.

[381] Höglund K, Bogstedt A, Fabre S, Aziz A, Annas P, Basun H, et al. Longitudinal stability evaluation of biomarkers and their correlation in cerebrospinal fluid and plasma from patients with Alzheimer's disease. *Journal of Alzheimer's Disease*. 2012;32(4):939–947.

[382] Li J, Llano DA, Ellis T, LeBlond D, Bhathena A, Jhee SS, et al. Effect of human cerebrospinal fluid sampling frequency on amyloid-β levels. *Alzheimer's & Dementia*. 2012 Jul;8(4):295–303.

[383] Le Bastard N, Aerts L, Sleegers K, Martin JJ, Van Broeckhoven C, De Deyn PP, et al. Longitudinal Stability of Cerebrospinal Fluid Biomarker Levels: Fulfilled Requirement for Pharmacodynamic Markers in Alzheimer's Disease. *Journal of Alzheimer's Disease*. 2012 Oct;.

[384] Helsel DR. Nondetects and data analysis. Statistics for censored environmental data. Wiley-Blackwell; 2005.

[385] Hewett P, Ganser GH. A comparison of several methods for analyzing censored data. *The Annals of occupational hygiene*. 2007 Oct;51(7):611–632.

[386] Uh HW, Hartgers FC, Yazdanbakhsh M, Houwing-Duistermaat JJ. Evaluation of regression methods when immunological measurements are constrained by detection limits. *BMC Immunology*. 2008;9(1):59.

[387] Helsel D. Much Ado About Next to Nothing: Incorporating Nondetects in Science. *Annals of Occupational Hygiene*. 2010 Mar;54(3):257–262.

[388] Helsel DR. Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere*. 2006 Dec;65(11):2434–2439.

[389] Ganser GH, Hewett P. An accurate substitution method for analyzing censored data. *Journal of Occupational and Environmental Hygiene*. 2010 Apr;7(4):233–244.

[390] Eriksson L, Johansson E, Kettaneh-Wold N, Trygg J, Wikström C, Wold S. *Multi- and Megavariate Data Analysis. Part I Basic Principles and Applications*. 2nd ed. Umea: MKS Umetrics AB; 2006.

[391] Swarbrick B. Multivariate Data Analysis for Dummies. Chichester: John Wiley & Sons; 2012.

[392] Wold S, Sjöström M, Eriksson L. PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*. 2001;58(2):109–130.

[393] van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*. 2006;7:142.

[394] Ahlbäck S. Osteoarthrosis of the knee. A radiographic investigation. *Acta Radiologica: Diagnosis*. 1968;p. Suppl 277:7–72.

[395] Barrett JP, Rashkoff E, Sirna EC, Wilson A. Correlation of roentgenographic patterns and clinical manifestations of symptomatic idiopathic osteoarthritis of the knee. *Clinical Orthopaedics and Related Research*. 1990 Apr;(253):179–183.

[396] Metcalfe AJ, Andersson MLE, Goodfellow R, Thorstensson CA. Is knee osteoarthritis a symmetrical disease? Analysis of a 12 year prospective cohort study. *BMC Musculoskeletal Disorders*. 2012;13:153.

[397] Keyes GW, Carr AJ, Miller RK, Goodfellow JW. The radiographic classification of medial gonarthrosis. Correlation with operation methods in 200 knees. *Acta Orthopaedica Scandinavica*. 1992 Oct;63(5):497–501.

[398] Gulati A, Chau R, Beard DJ, Price AJ, Gill HS, Murray DW. Localization of the full-thickness cartilage lesions in medial and lateral unicompartmental knee osteoarthritis. *Journal of Orthopaedic Research*. 2009 Oct;27(10):1339–1346.

[399] Rout R, Mcdonnell S, Hulley P, Jayadev C, Khan T, Carr A, et al. The pattern of cartilage damage in antero-medial osteoarthritis of the knee and its relationship to the anterior cruciate ligament. *Journal of Orthopaedic Research*. 2013 Feb;.

[400] Rout R, Mcdonnell S, Benson R, Athanasou N, Carr A, Doll H, et al. The histological features of anteromedial gonarthrosis–the comparison of two grading systems in a human phenotype of osteoarthritis. *The Knee*. 2011 Jun;18(3):172–176.

[401] Snelling S, Rout R, Davidson R, Clark I, Carr A, Hulley PA, et al. A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. *Osteoarthritis and Cartilage*. 2014 Feb;22(2):334–343.

[402] Rout R, Mcdonnell SM, Jayadev C, Xia Z, Murray D, Hulley PA, et al. Apoptosis and Reactive Oxygen Species – Possible Targets for Treatment of Osteoarthritis Shown in Anteromedial Gonarthrosis. *Journal of Bone and Joint Surgery (British)*. 2012;94-B(SUPP IX):16.

[403] Rout R, Mcdonnell SM, Hollander AP, Clark IM, Simms T, Davidson R, et al. Cartilage Changes in Anteromedial Osteoarthritis of the Knee - Increased Type I Collagen in Undamaged Cartilage. *Journal of Bone and Joint Surgery (British)*. 2009 Sep;91-B(SUPP III):468–469.

[404] Marascuilo LA. Large-sample multiple comparisons. *Psychological Bulletin*. 1966 May;65(5):280–290.

[405] Moen TC, Laskin W, Puri L. The lateral compartment in knees with isolated medial and patellofemoral osteoarthritis a histologic analysis of articular cartilage. *Journal of Arthroplasty*. 2011 Aug;26(5):783–787.

[406] Obeid EM, Adams MA, Newman JH. Mechanical properties of articular cartilage in knees with unicompartmental osteoarthritis. *Journal of Bone and Joint Surgery (British)*. 1994 Mar;76(2):315–319.

[407] Bert JM, Leverone J. Histologic appearance of "pristine" articular cartilage in knees with unicompartmental osteoarthritis. *Journal of Knee Surgery*. 2007 Jan;20(1):15–19.

[408] Brocklehurst R, Bayliss MT, Maroudas A, Coysh HL, Freeman MA, Revell PA, et al. The composition of normal and osteoarthritic articular cartilage from human knee joints. With special reference to unicompartmental replacement and osteotomy of the knee. *Journal of Bone and Joint Surgery (American)*. 1984 Jan;66(1):95–106.

[409] Wallis WJ, Simkin PA, Nelp WB, Foster DM. Intraarticular volume and clearance in human synovial effusions. *Arthritis and Rheumatism*. 1985 Apr;28(4):441–449.

[410] Myers SL, O'Connor BL, Brandt KD. Accelerated clearance of albumin from the osteoarthritic knee: Implications for interpretation of concentrations of 'cartilage markers' in synovial fluid. *Journal of Rheumatology*. 1996;23(10):1744–1748.

[411] Myers SL, Brandt KD, Eilam O. Even low-grade synovitis significantly accelerates the clearance of protein from the canine knee: Implications for Measurement of synovial fluid "markers" of osteoarthritis. *Arthritis and Rheumatism*. 1995;38(8):1085–1091.

[412] Loeuille D, Chary-Valckenaere I, Champigneulle J, Rat AC, Toussaint F, Pinzano-Watrin A, et al. Macroscopic and microscopic features of synovial membrane inflammation in the osteoarthritic knee: correlating magnetic resonance imaging findings with disease severity. *Arthritis and Rheumatism*. 2005 Nov;52(11):3492–3501.

[413] Samuels J, Krasnokutsky S, Abramson SB. Osteoarthritis: a tale of three tissues. *Bulletin of the NYU Hospital for Joint Diseases*. 2008;66(3):244–250.

[414] Feldmann M, Brennan FM, Maini RN. Role of cytokines in rheumatoid arthritis. *Annual Review of Immunology*. 1996;14:397–440.

[415] McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. *Nature Reviews Immunology*. 2007 Jun;7(6):429–442.

[416] Schlaak JF, Pfers I, Meyer Zum Büschenfelde KH, Märker-Hermann E. Different cytokine profiles in the synovial fluid of patients with osteoarthritis, rheumatoid arthritis and seronegative spondylarthropathies. *Clinical and Experimental Rheumatology*. 1996 Mar;14(2):155–162.

[417] Kokebie R, Aggarwal R, Lidder S, Hakimiyan AA, Rueger DC, Block JA, et al. The role of synovial fluid markers of catabolism and anabolism in osteoarthritis, rheumatoid arthritis and asymptomatic organ donors. *Arthritis Research and Therapy*. 2011 Mar;13(2):R50.

[418] Porée B, Kypriotou M, Chadjichristos C, Beauchef G, Renard E, Legendre F, et al. Interleukin-6 (IL-6) and/or soluble IL-6 receptor down-regulation of human type II collagen gene expression in articular chondrocytes requires a decrease of

Sp1.Sp3 ratio and of the binding activity of both factors to the COL2A1 promoter. *Journal of Biological Chemistry*. 2008 Feb;283(8):4850–4865.

[419] Scanzello CR, Umoh E, Pessler F, Diaz-Torne C, Miles T, Dicarlo E, et al. Local cytokine profiles in knee osteoarthritis: elevated synovial fluid interleukin-15 differentiates early from end-stage disease. *Osteoarthritis and Cartilage*. 2009 Aug;17(8):1040–1048.

[420] Richette P, François M, Vicaut E, Fitting C, Bardin T, Corvol M, et al. A high interleukin 1 receptor antagonist/IL-1beta ratio occurs naturally in knee osteoarthritis. *The Journal of Rheumatology*. 2008 Aug;35(8):1650–1654.

[421] Daheshia M, Yao JQ. The interleukin 1beta pathway in the pathogenesis of osteoarthritis. *The Journal of Rheumatology*. 2008 Dec;35(12):2306–2312.

[422] Kim WJ, Kang YJ, Koh EM, Ahn KS, Cha HS, Lee WH. LIGHT is involved in the pathogenesis of rheumatoid arthritis by inducing the expression of pro-inflammatory cytokines and MMP-9 in macrophages. *Immunology*. 2005 Feb;114(2):272–279.

[423] Ishida S, Yamane S, Ochi T, Nakano S, Mori T, Juji T, et al. LIGHT induces cell proliferation and inflammatory responses of rheumatoid arthritis synovial fibroblasts via lymphotoxin beta receptor. *The Journal of Rheumatology*. 2008 Jun;35(6):960–968.

[424] Pierer M, Brentano F, Rethage J, Wagner U, Hantzschel H, Gay RE, et al. The TNF superfamily member LIGHT contributes to survival and activation of synovial fibroblasts in rheumatoid arthritis. *Rheumatology*. 2007 Jul;46(7):1063–1070.

[425] Hayashi S, Miura Y, Tateishi K, Takahashi M, Kurosaka M. Decoy receptor 3 is highly expressed in patients with rheumatoid arthritis. *Modern Rheumatology*. 2010;20(1):63–68.

[426] Schulze-Tanzil G, Zreiqat H, Sabat R, Kohl B, Halder A, Müller RD, et al. Interleukin-10 and articular cartilage: experimental therapeutical approaches in cartilage disorders. *Current Gene Therapy*. 2009 Aug;9(4):306–315.

[427] Hampel U, Sesselmann S, Iserovich P, Sel S, Paulsen F, Sack R. Chemokine and cytokine levels in osteoarthritis and rheumatoid arthritis synovial fluid. *Journal of Immunological Methods*. 2013 Sep;p. 1–6.

[428] Harris Q, Seto J, O'Brien K, Lee PS, Kondo C, Heard BJ, et al. Monocyte Chemotactic Protein-1 Inhibits Chondrogenesis of Synovial Mesenchymal Progenitor Cells: An in vitro study. *Stem Cells*. 2013 Oct;31(10):2253–2265.

[429] Heard BJ, Fritzler MJ, Wiley JP, McAllister J, Martin L, El-Gabalawy H, et al. Intraarticular and systemic inflammatory profiles may identify patients with osteoarthritis. *The Journal of Rheumatology*. 2013 Aug;40(8):1379–1387.

[430] Saetan N, Honsawek S, Tanavalee A, Tantavisut S, Yuktanandana P, Parkpian V. Association of plasma and synovial fluid interferon-γ inducible protein-10 with radiographic severity in knee osteoarthritis. *Clinical Biochemistry*. 2011;44(14-15):1218–1222.

[431] Ruschpler P, Lorenz P, Eichler W, Koczan D, Hänel C, Scholz R, et al. High CXCR3 expression in synovial mast cells associated with CXCL9 and CXCL10 expression in inflammatory synovial tissues of patients with rheumatoid arthritis. *Arthritis Research and Therapy*. 2003;5(5):R241–52.

[432] Antonelli A, Fallahi P, Delle Sedie A, Ferrari SM, Maccheroni M, Bombardieri S, et al. High values of Th1 (CXCL10) and Th2 (CCL2) chemokines in patients with psoriatic arthtritis. *Clinical and Experimental Rheumatology*. 2009 Jan;27(1):22–27.

[433] Kuan WP, Tam LS, Wong CK, Ko FWS, Li T, Zhu T, et al. CXCL 9 and CXCL 10 as Sensitive markers of disease activity in patients with rheumatoid arthritis. *The Journal of Rheumatology*. 2010 Feb;37(2):257–264.

[434] Hsu YH, Hsieh MS, Liang YC, Li CY, Sheu MT, Chou DT, et al. Production of the chemokine eotaxin-1 in osteoarthritis and its role in cartilage degradation. *Journal of Cellular Biochemistry*. 2004 Nov;93(5):929–939.

[435] Chao PZ, Hsieh MS, Cheng CW, Lin YF, Chen CH. Regulation of MMP-3 expression and secretion by the chemokine eotaxin-1 in human chondrocytes. *Journal of Biomedical Science*. 2011 Nov;18(1):86.

[436] Pierzchala AW, Kusz DJ, Hajduk G. CXCL8 and CCL5 expression in synovial fluid and blood serum in patients with osteoarthritis of the knee. *Archivum Immunologiae et Therapiae Experimentalis*. 2011 Apr;59(2):151–155.

[437] Morales TI, Joyce ME, Sobel ME, Danielpour D, Roberts AB. Transforming growth factor-beta in calf articular cartilage organ cultures: synthesis and distribution. *Archives of Biochemistry and Biophysics*. 1991 Aug;288(2):397–405.

[438] Moos V, Fickert S, Muller B, Weber U, Sieper J. Immunohistological analysis of cytokine expression in human osteoarthritic and healthy cartilage. *The Journal of Rheumatology*. 1999 Apr;26(4):870–879.

[439] Blaney Davidson EN, Vitters EL, van der Kraan PM, van den Berg WB. Expression of transforming growth factor- (TGF) and the TGF signalling molecule SMAD-2P in spontaneous and instability-induced osteoarthritis: role in cartilage degradation, chondrogenesis and osteophyte formation. *Annals of the Rheumatic Diseases*. 2006 Nov;65(11):1414–1421.

[440] van den Berg WB. Osteoarthritis year 2010 in review: pathomechanisms. *Osteoarthritis and Cartilage*. 2011 Apr;19(4):338–341.

[441] Finnson KW, Chi Y, Bou-Gharios G, Leask A, Philip A. TGF-beta signaling in cartilage homeostasis and osteoarthritis. *Frontiers in Bioscience*. 2012;S4(1):251–268.

[442] Burrage PS, Mix KS, Brinckerhoff CE. Matrix metalloproteinases: role in arthritis. *Frontiers in Bioscience*. 2006;11:529–543.

[443] Brennan FM, Chantry D, Turner M, Foxwell B, Maini R, Feldmann M. Detection of transforming growth factor-beta in rheumatoid arthritis synovial tissue: lack of effect on spontaneous cytokine production in joint cell cultures. *Clinical and Experimental Immunology*. 1990 Aug;81(2):278–285.

[444] Aigner T, Soeder S, Haag J. IL-1beta and BMPs–interactive players of cartilage matrix degradation and regeneration. *European Cells and Materials*. 2006;12:49–56; discussion 56.

[445] Chubinskaya S, Frank BS, Michalska M, Kumar B, Merrihew CA, Thonar EJMA, et al. Osteogenic protein 1 in synovial fluid from patients with rheumatoid arthritis or osteoarthritis: relationship with disease and levels of hyaluronan and antigenic keratan sulfate. *Arthritis Research and Therapy*. 2006;8(3):R73.

[446] Honsawek S, Chayanupatkul M, Tanavalee A, Sakdinakiattikoon M, Deepaisarnsakul B, Yuktanandana P, et al. Relationship of plasma and synovial fluid BMP-7 with disease severity in knee osteoarthritis patients: a pilot study. *International Orthopaedics*. 2009 Aug;33(4):1171–1175.

[447] Fukui N, Zhu Y, Maloney WJ, Clohisy J, Sandell LJ. Stimulation of BMP-2 expression by pro-inflammatory cytokines IL-1 and TNF-α in normal and osteoarthritic chondrocytes. *Journal of Bone and Joint Surgery (American)*. 2003;85(SUPPL. 3):59–66.

[448] Lories RJU, Derese I, Ceuppens JL, Luyten FP. Bone morphogenetic proteins 2 and 6, expressed in arthritic synovium, are regulated by proinflammatory cytokines and differentially modulate fibroblast-like synoviocyte apoptosis. *Arthritis and Rheumatism*. 2003 Oct;48(10):2807–2818.

[449] Fukui N, Ikeda Y, Ohnuki T, Hikita A, Tanaka S, Yamane S, et al. Pro-inflammatory cytokine tumor necrosis factor-α induces bone morphogenetic protein-2 in chondrocytes via mRNA stabilization and transcriptional up-regulation. *Journal of Biological Chemistry*. 2006;281(37):27229–27241.

[450] Nakase T, Miyaji T, Tomita T, Kaneko M, Kuriyama K, Myoui A, et al. Localization of bone morphogenetic protein-2 in human osteoarthritic cartilage and osteophyte. *Osteoarthritis and Cartilage*. 2003 Apr;11(4):278–284.

[451] Galasso O, Familiari F, De Gori M, Gasparini G. Recent Findings on the Role of Gelatinases (Matrix Metalloproteinase-2 and -9) in Osteoarthritis. *Advances in Orthopedics*. 2012;2012:1–7.

[452] Vignon E, Balblanc JC, Mathieu P, Louisot P, Richard M. Metalloprotease activity, phospholipase A2 activity and cytokine concentration in osteoarthritis synovial fluids. *Osteoarthritis and Cartilage*. 1993 Apr;1(2):115–120.

[453] Yoshihara Y, Nakamura H, Obata K, Yamada H, Hayakawa T, Fujikawa K, et al. Matrix metalloproteinases and tissue inhibitors of metalloproteinases in synovial fluids from patients with rheumatoid arthritis or osteoarthritis. *Annals of the Rheumatic Diseases*. 2000 Jun;59(6):455–461.

[454] Tchetverikov I, Lohmander LS, Verzijl N, Huizinga TWJ, TeKoppele JM, Hanemaaijer R, et al. MMP protein and activity levels in synovial fluid from patients with joint injury, inflammatory arthritis, and osteoarthritis. *Annals of the Rheumatic Diseases*. 2005 May;64(5):694–698.

[455] Kim KS, Choi HM, Lee YA, Choi IA, Lee SH, Hong SJ, et al. Expression levels and association of gelatinases MMP-2 and MMP-9 and collagenases MMP-1 and MMP-13 with VEGF in synovial fluid of patients with arthritis. *Rheumatology International*. 2011 Apr;31(4):543–547.

[456] Heard BJ, Martin L, Rattner JB, Frank CB, Hart DA, Krawetz R. Matrix metalloproteinase protein expression profiles cannot distinguish between normal and early osteoarthritic synovial fluid. *BMC Musculoskeletal Disorders*. 2012;13:126.

[457] Ryu JH, Lee A, Huh MS, Chu J, Kim K, Kim BS, et al. Measurement of MMP Activity in Synovial Fluid in Cases of Osteoarthritis and Acute Inflammatory Conditions of the Knee Joints Using a Fluorogenic Peptide Probe-Immobilized Diagnostic Kit. *Theranostics*. 2012;2(2):198–206.

[458] Marini S, Fasciglione GF, Monteleone G, Maiotti M, Tarantino U, Coletta M. A correlation between knee cartilage degradation observed by arthroscopy and synovial proteinases activities. *Clinical Biochemistry*. 2003 Jun;36(4):295–304.

[459] Sandy JD, Verscharen C. Analysis of aggrecan in human knee cartilage and synovial fluid indicates that aggrecanase (ADAMTS) activity is responsible for the catabolic turnover and loss of whole aggrecan whereas other protease

activity is required for C-terminal processing in vivo. *The Biochemical Journal*. 2001;358(3):615–626.

[460] Struglics A, Larsson S, Pratta MA, Kumar S, Lark MW, Lohmander LS. Human osteoarthritis synovial fluid and joint cartilage contain both aggrecanase- and matrix metalloproteinase-generated aggrecan fragments. *Osteoarthritis and Cartilage*. 2006 Feb;14(2):101–113.

[461] Sandy JD. A contentious issue finds some clarity: on the independent and complementary roles of aggrecanase activity and MMP activity in human joint aggrecanolysis. *Osteoarthritis and Cartilage*. 2006 Feb;14(2):95–100.

[462] Durigova M, Nagase H, Mort JS, Roughley PJ. MMPs are less efficient than ADAMTS5 in cleaving aggrecan core protein. *Matrix Biology*. 2011 Mar;30(2):145–153.

[463] Struglics A, Hansson M. MMP proteolysis of the human extracellular matrix protein aggrecan is mainly a process of normal turnover. *The Biochemical Journal*. 2012 Sep;446(2):213–223.

[464] Ries C. Cytokine functions of TIMP-1. *Cellular and Molecular Life Sciences*. 2014 Feb;71(4):659–672.

[465] Aigner T, Zhu Y, Chansky HH, Matsen FA, Maloney WJ, Sandell LJ. Reexpression of type IIA procollagen by adult articular chondrocytes in osteoarthritic cartilage. *Arthritis and Rheumatism*. 1999 Jul;42(7):1443–1450.

[466] Salminen H, Vuorio E, Säämänen AM. Expression of Sox9 and type IIA procollagen during attempted repair of articular cartilage damage in a transgenic mouse model of osteoarthritis. *Arthritis and Rheumatism*. 2001 Apr;44(4):947–955.

[467] Rousseau JC, Zhu Y, Miossec P, Vignon E, Sandell LJ, Garnero P, et al. Serum levels of type IIA procollagen amino terminal propeptide (PIIANP) are decreased in patients with knee osteoarthritis and rheumatoid arthritis. *Osteoarthritis and Cartilage*. 2004 Jun;12(6):440–447.

[468] Garnero P, Ayral X, Rousseau JC, Christgau S, Sandell LJ, Dougados M, et al. Uncoupling of type II collagen synthesis and degradation predicts progression of joint damage in patients with knee osteoarthritis. *Arthritis and Rheumatism*. 2002 Oct;46(10):2613–2624.

[469] Sharif M, Kirwan J, Charni N, Sandell LJ, Whittles C, Garnero P. A 5-yr longitudinal study of type IIA collagen synthesis and total type II collagen degradation in patients with knee osteoarthritis–association with disease progression. *Rheumatology*. 2007 Jun;46(6):938–943.

[470] Berry PA, Maciewicz RA, Wluka AE, Downey-Jones MD, Forbes A, Hellawell CJ, et al. Relationship of serum markers of cartilage metabolism to imaging and clinical outcome measures of knee joint structure. *Annals of the Rheumatic Diseases*. 2010 Oct;69(10):1816–1822.

[471] Swärd P, Frobell R, Englund M, Roos H, Struglics A. Cartilage and bone markers and inflammatory cytokines are increased in synovial fluid in the acute phase of knee injury (hemarthrosis) - a cross-sectional analysis. *Osteoarthritis and Cartilage*. 2012 Nov;20(11):1302–1308.

[472] Larsson S, Lohmander LS, Struglics A. Synovial fluid level of aggrecan ARGS fragments is a more sensitive marker of joint disease than glycosaminoglycan or aggrecan levels: a cross-sectional study. *Arthritis Research and Therapy*. 2009;11(3):R92.

[473] Dahlberg L, Ryd L, Heinegård D, Lohmander LS. Proteoglycan fragments in joint fluid. Influence of arthrosis and inflammation. *Acta Orthopaedica Scandinavica*. 1992 Aug;63(4):417–423.

[474] El-Arman MM, El-Fayoumi G, El-Shal E, El-Boghdady I, El-Ghaweet A. Aggrecan and cartilage oligomeric matrix protein in serum and synovial fluid of patients with knee osteoarthritis. *HSS Journal*. 2010 Sep;6(2):171–176.

[475] Lohmander LS, Saxne T, Heinegård DK. Release of cartilage oligomeric matrix protein (COMP) into joint fluid after knee injury and in osteoarthritis. *Annals of the Rheumatic Diseases*. 1994 Jan;53(1):8–13.

[476] Di Cesare PE, Carlson CS, Stolerman ES, Hauser N, Tulli H, Paulsson M. Increased degradation and altered tissue distribution of cartilage oligomeric matrix protein in human rheumatoid and osteoarthritic cartilage. *Journal of Orthopaedic Research*. 1996 Nov;14(6):946–955.

[477] Neidhart M, Hauser N, Paulsson M, DiCesare PE, Michel BA, Häuselmann HJ. Small fragments of cartilage oligomeric matrix protein in synovial fluid and serum as markers for cartilage degradation. *British Journal of Rheumatology*. 1997 Nov;36(11):1151–1160.

[478] Ganu V, Goldberg R, Peppard J, Rediske J, Melton R, Hu SI, et al. Inhibition of interleukin-1alpha-induced cartilage oligomeric matrix protein degradation in bovine articular cartilage by matrix metalloproteinase inhibitors: potential role for matrix metalloproteinases in the generation of cartilage oligomeric matrix protein fragments in arthritic synovial fluid. *Arthritis and Rheumatism*. 1998 Dec;41(12):2143–2151.

[479] Dickinson SC, Vankemmelbeke MN, Buttle DJ, Rosenberg K, Heinegård D, Hollander AP. Cleavage of cartilage oligomeric matrix protein (thrombospondin-5) by matrix metalloproteinases and a disintegrin and metalloproteinase with thrombospondin motifs. *Matrix Biology*. 2003 May;22(3):267–278.

[480] Lai Y, Yu XP, Zhang Y, Tian Q, Song H, Mucignat MT, et al. Enhanced COMP catabolism detected in serum of patients with arthritis and animal disease models through a novel capture ELISA. *Osteoarthritis and Cartilage*. 2012 Aug;20(8):854–862.

[481] Crnkic M, Månsson B, Larsson L, Geborek P, Heinegård D, Saxne T. Serum cartilage oligomeric matrix protein (COMP) decreases in rheumatoid arthritis patients treated with infliximab or etanercept. *Arthritis Research and Therapy*. 2003;5(4):R181–185.

[482] Morozzi G, Fabbroni M, Bellisai F, Pucci G, Galeazzi M. Cartilage oligomeric matrix protein level in rheumatic diseases: potential use as a marker for measuring articular cartilage damage and/or the therapeutic efficacy of treatments. *Annals of the New York Academy of Sciences*. 2007 Jun;1108:398–407.

[483] Huebner JL, Kraus VB. Assessment of the utility of biomarkers of osteoarthritis in the guinea pig. *Osteoarthritis and Cartilage*. 2006 Sep;14(9):923–930.

[484] Rübenhagen R, Schüttrumpf JP, Stürmer KM, Frosch KH. Interleukin-7 levels in synovial fluid increase with age and MMP-1 levels decrease with progression of osteoarthritis. *Acta Orthopaedica*. 2011 Dec;.

[485] Otterness IG, Swindell AC, Zimmerer RO, Poole AR, Ionescu M, Weiner E. An analysis of 14 molecular markers for monitoring osteoarthritis: segregation of the markers into clusters and distinguishing osteoarthritis at baseline. *Osteoarthritis and Cartilage*. 2000 May;8(3):180–185.

[486] Otterness IG, Weiner E, Swindell AC, Zimmerer RO, Ionescu M, Poole AR. An analysis of 14 molecular markers for monitoring osteoarthritis. Relationship of the markers to clinical end-points. *Osteoarthritis and Cartilage*. 2001 Apr;9(3):224–231.

[487] Anitua E, Sánchez M, de La Fuente M, Azofra J, Zalduendo M, Aguirre JJ, et al. Relationship between Investigative Biomarkers and Radiographic Grading in Patients with Knee Osteoarthritis. *International Journal of Rheumatology*. 2009;2009:747432.

[488] Raza K, Falciani F, Curnow S, Ross E, Lee C, Akbar A, et al. Early rheumatoid arthritis is characterized by a distinct and transient synovial fluid cytokine profile of T cell and stromal cell origin. *Arthritis Research and Therapy*. 2005 Jan;7(4):R784–795.

[489] Pan X, Huang L, Chen J, Dai Y, Chen X. Analysis of synovial fluid in knee joint of osteoarthritis:5 proteome patterns of joint inflammation based on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *International Orthopaedics*. 2012 Jan;36(1):57–64.

[490] Gobezie R, Kho A, Krastins B, Sarracino DA, Thornhill TS, Chase M, et al. High abundance synovial fluid proteome: distinct profiles in health and osteoarthritis. *Arthritis Research and Therapy*. 2007;9(2):R36.

[491] Afara I, Prasadam I, Crawford R, Xiao Y, Oloyede A. Non-destructive evaluation of articular cartilage defects using near-infrared (NIR) spectroscopy in osteoarthritic rat models and its direct relation to Mankin score. *Osteoarthritis and Cartilage*. 2012 Nov;20(11):1367–1373.

[492] Afara I, Singh S, Oloyede A. Application of near infrared (NIR) spectroscopy for determining the thickness of articular cartilage. *Medical Engineering & Physics*. 2013 Jan;35(1):88–95.

[493] Hanifi A, Bi X, Yang X, Kavukcuoglu B, Lin PC, DiCarlo E, et al. Infrared fiber optic probe evaluation of degenerative cartilage correlates to histological grading. *The American Journal of Sports Medicine*. 2012 Dec;40(12):2853–2861.

[494] Li G, Thomson M, DiCarlo E, Yang X, Nestor B, Bostrom MPG, et al. A chemometric analysis for evaluation of early-stage cartilage degradation by infrared fiber-optic probe spectroscopy. *Applied Spectroscopy*. 2005 Dec;59(12):1527–1533.

[495] Lamers RJAN, van Nesselrooij JHJ, Kraus VB, Jordan JM, Renner JB, Dragomir AD, et al. Identification of an urinary metabolite profile associated with osteoarthritis. *Osteoarthritis and Cartilage*. 2005 Sep;13(9):762–768.

[496] Lamers RJAN, DeGroot J, Spies-Faber EJ, Jellema RH, Kraus VB, Verzijl N, et al. Identification of disease- and nutrient-related metabolic fingerprints in osteoarthritic Guinea pigs. *Journal of Nutrition*. 2003 Jun;133(6):1776–1780.

[497] Gierman LM, Wopereis S, van El B, Verheij ER, Werff-van der Vat BJC, Bastiaansen-Jenniskens YM, et al. Metabolic profiling reveals differences in concentrations of oxylipins and fatty acids secreted by the infrapatellar fat pad of end-stage osteoarthritis and normal donors. *Arthritis and Rheumatism*. 2013 Jul;.

[498] Jayadev C, Khan T, Coulter A, Beard DJ, Price AJ. Patient decision aids in knee replacement surgery. *The Knee*. 2012 Mar;.

[499] Felson DT. Osteoarthritis: Virtual joint replacement as an outcome measure in OA. *Nature Reviews Rheumatology*. 2012 Apr;8(4):187–188.

[500] Shrivastava A, Gupta V. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chronicles of Young Scientists*. 2011;2(1):21.

[501] Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist Reviews*. 2008 Aug;29 Suppl 1:S49–52.

[502] Anderson DJ. Determination of the lower limit of detection. *Clinical Chemistry*. 1989 Oct;35(10):2152–2153.

[503] Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, et al. Fit-for-purpose method development and validation for successful biomarker measurement. *Pharmaceutical Research*. 2006 Feb;23(2):312–328.

[504] Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*. 2005 Mar;85(3):257–268.