

***Methodological development of molecular endotype discovery from synovial fluid of individuals with knee osteoarthritis: the STEpUP OA Consortium***

**Authors:** Y.Deng<sup>1\*</sup>, T.A.Perry<sup>1\*</sup>, P.Hulley<sup>2</sup>, R.A.Maciewicz<sup>1</sup>, J.Mitchelmore<sup>3</sup>, D.Perry<sup>4</sup>, S.Larsson<sup>5</sup>, S. Brachat<sup>3</sup>, A.Struglics<sup>5</sup>, C.T. Appleton<sup>6</sup>, S. Kluzek<sup>2,7</sup>, N. K. Arden<sup>2,8</sup>, D. Felson<sup>9</sup>, B.Marsden<sup>2,10</sup>, B.D.M.Tom<sup>11</sup>, L.Bondi<sup>11</sup>, M. Kapoor<sup>12</sup>, V.Batchelor<sup>1</sup>, J. Mackay-Alderson<sup>1</sup>, V.Kumar<sup>2</sup>, L. S. Lohmander<sup>5</sup>, T. J. Welting<sup>13</sup>, D. A. Walsh<sup>14,15</sup>, A.M.Valdes<sup>14</sup>, the STEpUP OA Consortium, T. L. Vincent<sup>1#</sup>, F. E. Watt<sup>16</sup>, L. Jostins-Dean<sup>1#</sup>

<sup>1</sup> Centre for Osteoarthritis Pathogenesis Versus Arthritis, Kennedy Institute of Rheumatology, NDORMS, University of Oxford, Oxford, UK.

<sup>2</sup> Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, UK.

<sup>3</sup> Novartis Institutes for Biomedical Research, Basel, Switzerland.

<sup>4</sup> SomaLogic, Boulder, Colorado, USA.

<sup>5</sup> Faculty of Medicine, Department of Clinical Sciences Lund, Orthopaedics, Lund University, Lund, Sweden.

<sup>6</sup> Bone and Joint Institute, University of Western Ontario, London, Ontario, Canada.

<sup>7</sup> NIHR Nottingham Biomedical Research Centre and Versus Arthritis Sport, Exercise and Osteoarthritis Centre, University of Nottingham, Nottingham, UK.

<sup>8</sup> Centre for Sport, Exercise and Osteoarthritis Research Versus Arthritis, University of Oxford, Oxford, UK.

<sup>9</sup> Section of Rheumatology, Boston University School of Medicine, Boston, Massachusetts, USA.

<sup>10</sup> Nuffield Department of Medicine, University of Oxford, Oxford, UK.

<sup>11</sup> MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.

<sup>12</sup> Schroeder Arthritis Institute, University Health Network, Toronto, Ontario, Canada.

<sup>13</sup> Laboratory for Experimental Orthopedics, Department of Orthopedic Surgery, Maastricht University, Maastricht, Netherlands.

<sup>14</sup> Pain Centre Versus Arthritis, Advanced Pain Discovery Platform, and the NIHR Nottingham Biomedical Research Centre, University of Nottingham, Nottingham, UK.

<sup>15</sup> Sherwood Forest Hospitals NHS Foundation Trust, Sutton in Ashfield, UK.

<sup>16</sup> Department of Immunology and Inflammation, Imperial College London, London, UK.

\*Joint first authors, #equal contributions.

**Corresponding author:** Professor Tonia L Vincent (tonia.vincent@kennedy.ox.ac.uk), Dr Luke Jostins-Dean (luke.jostins@kennedy.ox.ac.uk) and Dr Fiona Watt (fiona.watt@kennedy.ox.ac.uk), Centre for Osteoarthritis Pathogenesis Versus Arthritis, Kennedy Institute of Rheumatology, University of Oxford, Oxford OX3 7FY, UK.

**Keywords:** biomarker, proteomic, synovial fluid, osteoarthritis, knee injury, endotype, stratification, SomaLogic.

**Running title:** Methods for molecular endotype discovery in STEpUP OA

## ABSTRACT

**Objectives:** To develop and validate a pipeline for quality controlled (QC) protein data for largescale analysis of synovial fluid (SF), using SomaLogic technology.

**Design:** Knee SF and associated clinical data were from partner cohorts. SF samples were centrifuged, supernatants stored at  $-80^{\circ}\text{C}$ , then analysed by SomaScan Discovery Plex V4.1 (>7000 SOMAmers/proteins).

**Setting:** An international consortium of 9 academic and 8 commercial partners (STEpUP OA).

**Participants:** 1746 SF samples from 1650 individuals comprising OA, joint injury, healthy controls and inflammatory arthritis controls, divided into discovery (n=1045) and replication (n=701) datasets.

**Primary and secondary outcome measures:** An optimised approach to standardisation was developed iteratively, monitoring reliability and precision (comparing coefficient of variation [%CV] of 'pooled' SF samples between plates and correlation with prior immunoassay for 9 analytes). Pre-defined technical confounders were adjusted for (by Limma) and batch correction was by ComBat. Poorly performing SOMAmers and samples were filtered. Variance in the data was determined by principal component (PC) analysis. Data were visualised by Uniform Manifold Approximation and Projection (UMAP).

**Results:** Optimal SF standardisation aligned with that used for plasma, but without median normalisation. There was good reliability (<20 %CV for >80% of SOMAmers in pooled samples) and overall good correlation with immunoassay. PC1 accounted for 48% of variance and strongly correlated with individual SOMAmer signal intensities (median correlation coefficient 0.70). These could be adjusted using an 'intracellular protein score'. PC2 (7% variance) was attributable to processing batch and was batch-corrected by ComBat. Lesser effects were attributed to other technical confounders. Data visualisation by UMAP revealed clustering of injury and OA cases in overlapping but distinguishable areas of high-dimensional proteomic space.

**Conclusions:** We define a standardised approach for SF analysis using the SOMAscan platform and identify likely 'intracellular' protein as being a major driver of variance in the data.

### Strengths and limitations:

- This is the largest number of individual synovial fluid samples analysed by a high content proteomic platform (SomaLogic technology)
- SomaScan offers reliable, precise relative SF data following standardisation for over 6000 proteins
- Significant variance in the data was driven by a protein signal which is likely intracellular in origin: it is not yet clear whether this is due to technical considerations, normal cell turnover or relevant pathological processes
- Adjusting for confounding factors might conceal the true structure of the data and reduce the ability to detect 'molecular endotypes' within disease groups

**Word**

**count:**

**5,708**

## 1 INTRODUCTION

2

3 Osteoarthritis (OA) is a highly prevalent and disabling condition and arguably represents one of the  
4 greatest unmet clinical needs of all musculoskeletal conditions<sup>[1, 2]</sup>, now recognised by the FDA as a  
5 ‘serious disease’<sup>[3, 4]</sup>. OA is a disease of the synovial joints manifesting as localised, low-grade  
6 inflammation of the synovium, cartilage damage and subchondral bone remodeling<sup>[5]</sup>, which lead to  
7 pain, stiffness and loss of function<sup>[6, 7]</sup>. Despite growing clinical demand and best efforts in pre-  
8 clinical models and translational studies to understand the underlying pathogenesis, target discovery  
9 and drug development for knee OA in humans have been slow<sup>[8]</sup>. Results from randomised clinical  
10 trials of putative disease-modifying osteoarthritis drugs (DMOADs) have been largely disappointing<sup>[9,</sup>  
11 <sup>10]</sup> with a few treatments showing modest effects on cartilage preservation<sup>[11, 12]</sup>.

12

13 OA might not be a single disease<sup>[13, 14]</sup> but rather a group of diseases with a similar clinical  
14 presentation but driven by distinct molecular pathways known as ‘endotypes’. These might  
15 determine the course of disease and in some cases predict response to treatment. It is presumed  
16 that molecular endotypes might relate to discernible patient characteristics and may help to explain  
17 the heterogeneity of OA ‘clinical phenotypes’<sup>[15, 16]</sup>. Many cellular processes have been proposed as  
18 critical drivers in OA pathogenesis such as immune-mediated inflammation<sup>[17]</sup>, mechanically-  
19 mediated inflammation (‘mechanoflammation’)<sup>[18]</sup>, low/failed tissue repair<sup>[19, 20]</sup> and cellular  
20 senescence<sup>[21]</sup>. These in turn may relate to a broad range of aetiological factors that are associated  
21 with OA<sup>[22-26]</sup>. Efforts have been made to classify subgroups of people with OA based on  
22 epidemiological factors<sup>[27]</sup>, with several clinically defined phenotypes now suggested in the  
23 literature<sup>[9, 28-30]</sup>. A recent systematic review of 24 studies reported that up to 84% of people with  
24 OA could be assigned to at least one of six phenotypes<sup>[28, 31]</sup>. These clinical phenotypes are, however,  
25 not mutually exclusive, and are poor systematic classifiers because they are a mixture of overlapping  
26 demographic, clinical, radiographic, aetiological and systemic features. They currently have limited  
27 clinical applicability<sup>[29]</sup> and there is a paucity of data relating them to distinct molecular pathways or  
28 to clinical outcomes in OA<sup>[32]</sup>.

29

30 Whilst there has been a plethora of studies of candidate molecules trying to identify diagnostic or  
31 prognostic biomarkers of OA, relatively few have used hypothesis-free approaches in large numbers  
32 of human biological samples to identify molecular endotypes<sup>[28, 33, 34]</sup>. Such collaborations are  
33 required to help move biomarker discovery forward<sup>[35]</sup>. Two broad matrices have been studied:  
34 blood (plasma or serum), which has the advantage of accessibility, and synovial fluid (SF), acquired  
35 by joint aspiration. SF has several advantages over blood for exploring molecular mechanisms.

36 Firstly, it has adjacency to joint tissues, and may reflect activities in synovium, bone as well as  
37 cartilage<sup>[36-38]</sup>. Secondly, concentrations of analytes within the SF provide an indication of biological  
38 activity and target tissue activation<sup>[39, 40]</sup>. Thirdly, the SF from a given joint is less confounded by  
39 disease at other sites than is, for example, blood. Finally, a number of analytes that are highly  
40 regulated in the SF are not reflected in the plasma<sup>[36, 41, 42]</sup>.

41

42 The Synovial fluid To detect molecular Endotypes by Unbiased Proteomics in OA (STEpUP OA)  
43 Consortium was set up to address a primary objective: to determine whether there are detectable  
44 distinct molecular endotypes in knee OA, through a hypothesis-free, unsupervised proteomic  
45 analysis applying SomaLogic array technology<sup>[43]</sup> of SF from a large number of participants with, or at  
46 increased risk of, knee OA. SomaScan, an aptamer-based proteomics technology, offers the ability to  
47 measure large numbers of protein analytes from a small volume of biological fluid. However,  
48 detailed methodology is lacking for quality control (QC) and data analysis pipelines specifically  
49 tailored to SF.

50 SF presents analysis challenges due to its complex matrix which is rich in hyaluronan making the fluid  
51 viscous, variability in joint effusion volume between and within patients, and potential  
52 contamination with blood at time of aspiration. To combat some of these challenges, hyaluronidase  
53 treatment of the fluid post aspiration or lavage of the joint prior to aspiration have been utilised<sup>[42,</sup>  
54 <sup>44]</sup>.

55

56 In this study we describe the processing and analysis of SF, and the optimisation of a standardised  
57 quality control (QC) and analysis pipeline for these data. We evaluate performance of SF on the  
58 SomaLogic platform at scale for the first time and identify important technical confounders requiring  
59 adjustment prior to downstream analysis. Prespecified potential confounding factors included those  
60 relating to sample processing or to the sample itself, such as its age, number of freeze-thaws, visible  
61 blood staining and sample volume. These investigations were used to inform STEpUP OA's primary  
62 data analysis plan (<https://www.kennedy.ox.ac.uk/oacentre/stepup-oa/stepup-oa>) and make our  
63 work replicable by others.

64

65

## 66 **METHODS**

67 Details of consortium structure, governance and ethical approvals can be found in Supplementary  
68 Methods. Working groups oversaw key activities (Figure S1). Six participating sites with 17  
69 participant collections (henceforth referred to as ‘cohorts’) including those with either knee OA or  
70 acute knee joint injury provided associated SF samples. Each had ethical approval (Table S1). In  
71 addition, the University of Oxford Medical Sciences Central University Research Ethics Committee  
72 (CUREC) granted ethical approval for the processing, storage and use of samples and linked data for  
73 this project on 1<sup>st</sup> November 2019 (R67029/RE001).

74

### 75 **Participant eligibility criteria**

76 All but one cohort had existing associated stored participant SF samples. Inclusion criteria were: i)  
77 evidence of a confirmed diagnosis of knee OA, or history of recent knee injury, ii) associated basic  
78 clinical information including (as a minimum) age at sampling, sex and indication of OA disease  
79 status, iii) a minimum volume of SF (90 µl, ideally 200 µl) and iv) SF had been centrifuged between  
80 1800-3000g, prior to supernatant storage at -80°C. Exclusion criteria were: i) additional forms of  
81 arthritis e.g. gout, rheumatoid arthritis, psoriatic arthritis, as determined by host investigator; ii)  
82 confounding medical conditions e.g. concurrent infection, cancer; iii) confounding treatments e.g.  
83 index knee surgery in the preceding 6 months, index knee steroid injection in preceding 3 months;  
84 iv) chemotherapy and; v) significant deviation in storage procedure (e.g. freezer drop-out defined by  
85 host investigator).

86

### 87 **Sample processing and SomaLogic assay**

88 Consortium samples: 1746 SF samples were eligible and processed for STEpUP OA. A STEpUP  
89 participant ID number (PIN) and related unique sample identification number (SIN) were generated  
90 for each participant and their associated sample(s). Sample processing was performed in Oxford in  
91 four tranches over a 24-month period. For analysis by SomaLogic, SF enzymatic digestion, using  
92 hyaluronidase, was carried out (Supplementary methods). Briefly, sufficient bovine testicular  
93 hyaluronidase (4mg/ml; Sigma-Aldrich) for the entire project was reconstituted as a single batch  
94 from a single lot number and frozen in aliquots until use. For each tranche, batch processing was  
95 performed over consecutive working days, i.e. over as short a time as possible. Briefly, a batch of SFs  
96 was thawed, centrifuged at 3000g at 20°C for 25 minutes and 175 µl of SF supernatant diluted 1:2  
97 with the same volume of hyaluronidase solution and agitated at room temperature for 1 hour,  
98 followed by further centrifugation for 5 minutes<sup>[42]</sup>. Supernatants were aliquoted and stored at -80°C

99 and transferred on dry ice by temperature-controlled shipping to SomaLogic (single shipment per  
100 tranche).

101

102 Consortium controls/QC samples: Equal volumes of SF samples from 6 participants per group were  
103 used to generate single batches of hyaluronidase-treated 'pooled samples' for each of OA and knee  
104 injury at the start of project. Subaliquots of these then acted as internal QC controls, being run on  
105 each SomaLogic plate, enabling calculation of intra-assay and inter-assay coefficients of variation  
106 (CVs), as well as assessing effects of freeze thaw (a multiple freeze-thawed aliquot), frozen storage  
107 of hyaluronidase (an untreated aliquot freshly treated with frozen hyaluronidase during processing  
108 of each tranche) and centrifugation (an additional unspun pooled sample, from 6 unspun OA SF  
109 samples). A further 18 samples (split at the time of collection, with the paired aliquot remaining  
110 'unspun') were included to further examine the effects of centrifugation. 42 other 'comparator'  
111 samples were included (disease-free controls from non-painful knees or from normal joints at  
112 amputation/post-mortem; samples from individuals with definite inflammatory arthritis). Three  
113 samples from three separate participants were re-processed under 3 different temperature  
114 conditions and re-analysed to examine the effects of laboratory re-processing. A subgroup of the  
115 freshly collected samples were processed specifically to test generalizability to OA SF which had not  
116 been centrifuged ('unspun') (n=18). 235 unspun OA samples were subsequently included in the  
117 replication analysis.

118

119 Samples were assayed on the SomaLogic SomaScan Discovery Plex V4.1 by SomaLogic, in Boulder,  
120 US. All samples from all 4 tranches were processed as a single batch on twenty-two sequential 96-  
121 well plates in January 2022. All samples were randomised within and between plates whilst ensuring  
122 appropriate controls on each plate. Each plate included 83 participant SF single samples; one pooled  
123 OA sample; one pooled knee injury sample; five plasma calibrator samples; three plasma QC samples  
124 and three blanks per plate. The SomaScan platform quantified 7,596 synthetic DNA SOMAmers (for  
125 7289 human targets) (Slow Off-rate Modified Aptamers<sup>[45, 46]</sup>) that bound to 6596 unique human  
126 proteins. The generated SomaScan protein quantification was securely transferred from SomaLogic  
127 to Oxford as .adat files.

128

129 Additional sample metadata, both from collecting sites (where available) and those generated in  
130 Oxford and at SomaLogic, included: sample blood staining (by visual staining defined by host  
131 investigator at time of collection); initial centrifugation of the sample; number of previous freeze  
132 thaws; the date of laboratory processing; batch and order of processing; the plate and position of



133 the sample. These sample metadata were defined as technical confounders in our QC pipeline (Table  
134 S2).

135

### 136 **Clinical data**

137 Pseudonymised associated participant clinical data were transferred from participating sites to  
138 Oxford, linked to their consortium PIN, mapped to variables where necessary and uploaded to a  
139 REDCap database (Research Electronic Data Capture, Vanderbilt University, US)<sup>[47]</sup>, hosted by the  
140 University of Oxford. Data integrity and completeness were ensured using a data dictionary, data  
141 entry constraints and a combination of automated, systematic and random checks by two of the  
142 study team.

143

144 A consortium working group oversaw all aspects of data management including definition of  
145 variables and associated data dictionary, data harmonisation and design of the database (Figure S1).  
146 Informed by their relative clinical importance, by data availability and by iterative review, a core  
147 clinical dataset (a subset of the data dictionary) was defined: the first phenotype release “Pheno 1”  
148 (demographic data and harmonised measures of radiographic disease severity) and the second  
149 release “Pheno 2” (dichotomous and continuous harmonised patient-reported outcome measures  
150 for knee pain<sup>[48]</sup>). (Table S2 & Supplementary methods).

151

### 152 **Data QC approach**

153 Methods to develop QC and data analysis pipelines prior to the primary discovery analysis were pre-  
154 defined in the Quality Assurance plan ([https://www.kennedy.ox.ac.uk/oacentre/stepup-oa/stepup-  
155 oa](https://www.kennedy.ox.ac.uk/oacentre/stepup-oa/stepup-oa)). This QC pipeline aimed to validate methods for standardisation of the data, through a series of  
156 normalisation steps (given that SF was a non-standard matrix on SomaScan), correction for technical  
157 confounders and filtering based on pre-defined quality thresholds for SOMAers, proteins and  
158 samples. The approach included pre-defined data exploration, though where issues were found,  
159 these were iteratively investigated and findings used to refine the QC pipeline. This approach was  
160 informed by our prior published work<sup>[42]</sup>, SomaLogic expertise, initial consortium pilot work on 435  
161 samples previously assessed on an earlier version (4.0) of the SomaScan platform, and subsequent  
162 QC work within this dataset.

163

164 The usual SomaScan analysis pipeline for plasma involves a series of standardisation procedures to  
165 reduce nuisance variance, using plasma calibrator and plasma QC samples included on plates to  
166 reduce the effect of technical factors across samples and plates<sup>[49]</sup>. This routine standardisation of

167 the SomaScan relative fluorescence units (RFU), adjusted by the protein's dilution factor used in the  
168 SomaScan assay (the "dilution bin"), was applied in a stepwise way, using i) hybridisation control  
169 normalisation (to remove well-to-well variation due to different rates of hybridisation between  
170 SOMAmers and fluorescence probes using spiked-in control SOMAmers); ii) plate scaling using  
171 plasma calibrators (to remove variation in overall intensity between plates); iii) median signal  
172 normalisation (to decrease variation due to total fluorescence intensity between samples); and iv)  
173 calibration (using plasma calibrator samples of known concentration to rescale each protein and  
174 reduce assay differences between runs). Each normalisation step was tested in a sequential manner.

175

176 After optimised standardisation, the R package *limma* was used to adjust proteins for a number of  
177 pre-specified continuous covariates, as described below. The batch correction method ComBat (in  
178 *sva* R package) was applied<sup>[50]</sup> to adjust the mean and variance of each protein for batch effects in  
179 datasets where the batch covariate, such as plate, is known<sup>[44]</sup> (again, as described below).

180

181 Based on our initial QC assessments, to quantify and adjust for differences in the contribution of  
182 intracellular proteins to the proteome, we defined an Intracellular Protein Score (IPS) for each  
183 sample  $i$ , as a weighted sum of protein concentrations using the equation

184

$$IPS_i = \sum_p d_p C_{ip}$$

185 where  $d_p$  is the Cohen's  $d$  for the difference in concentration for protein  $p$  between paired spun and  
186 unspun samples, and  $C_{ip}$  is the log concentration of protein  $p$  in sample  $i$ .

187

188 Checks of assay performance and biological validity were carried out by measuring repeatability  
189 using the pooled samples on each plate, use of metadata for prespecified technical confounders, and  
190 comparison with previously generated quantitative immunoassay data (R&D or Meso Scale  
191 Discovery, available for 60 OA and injury SF samples (without hyaluronidase treatment) for 9  
192 overlapping proteins (Table S3).

193

#### 194 **Statistics and Analysis**

195 Principal Components Analysis (PCA) was used to visualize proteome-wide patterns of variation in  
196 the data, with further visualisation of Principal Components (PCs) with Uniform Manifold  
197 Approximation and Projection (UMAP) 2-dimensional plots (UMAP applied to the set of top PCs that  
198 explained >80% of total variation). Various other bioinformatic, descriptive and statistical techniques  
199 were employed to test the quality of the data. We checked:

200

- 201 • Inter-assay repeatability – the %CV of each protein, i.e. the ratio of the standard deviation of the  
202 concentration to the mean of the concentration within repeated samples, and the proportion of  
203 non-technical variation ( $R^2$ ) for each protein, estimated as one minus the square of the ratio of  
204 the variance in repeated samples to the variance in non-repeated SF samples.
- 205 • The effect of freeze-thawing on normalised RFU signal (measured by %CV between repeatedly  
206 freeze-thawed and non-freeze-thawed samples).
- 207 • The effect of centrifugation on normalised RFU signal (by estimates of correlation between, and  
208 differential abundance of, proteins in unspun and spun samples, using correlation tests and  
209 paired t-tests respectively).
- 210 • Assay accuracy – comparing SomaScan normalised RFU signal with existing quantitative  
211 immunoassay data (by Pearson correlation coefficients).
- 212 • Effects of each technical confounder on standardised RFU signal (from combined cohorts). Linear  
213 regression analyses were applied to identify the most significant principal components (PCs) and  
214 proteins associated with technical confounders (Bonferonni adjusted  $p < 0.05$ ).

215

216 In addition, SOMAmers and samples of insufficient quality were removed as follows:

217

218 *SOMAmer filtering:*

- 219 • SOMAmers which were highly associated with pre-specified confounders (Bonferroni  
220 adjusted  $p < 0.05$ ) were removed (Table S4).
- 221 • SOMAmers from non-human organisms or control SOMAmers (including Spuriomer,  
222 hybridisation control elution, deprecated, non-biotin, and non-cleavable) were excluded.
- 223 • The estimated proportion of non-technical (i.e. biological) variation  $R^2$  was calculated using  
224 pooled SF samples (OA, knee injury), defined as  $(1 - V_{\text{Pooled}}/V_{\text{total}})^2$ , where  $V_{\text{Pooled}}$  and  $V_{\text{total}}$  are  
225 the variances ( $V$ ) in pooled and non-pooled (including all individual) samples respectively. If  
226  $R^2$  for a given SOMAmer accounted for less than 50% of total variation in either OA or knee  
227 injury it was removed.

228

229 *Sample filtering:*

- 230 • If a sample had more than 25% of protein values above or below the upper or lower  
231 limits of detection respectively, the sample was removed.

- 232
- 233
- 234
- 235
- 236
- 237
- 238
- 239
- 240
- 241
- 242
- 243
- 244
- 245
- 246
- 247
- (This was applying lower and upper limits of detection (LOD) (defined by SomaLogic), where lower LOD was:  $[median\ concentration\ of\ blanks] + 4.9 \times [median\ absolute\ deviation\ of\ blanks]$ , based on three blanks (i.e. buffer only) per plate. Upper LOD was defined as 80,000 RFU).
  - Identification of outliers by PCA: samples that were beyond 5 standard deviations (SDs) from the center of the principal component space (made up of PCs that explained at least 80% of variation) were removed.
  - Identification of total signal intensity outliers: samples that were beyond 5 SDs from the mean in total RFU distribution were removed. Total RFU of per sample was defined as the sum of all the RFU values of that sample.
  - Samples that were flagged by SomaLogic's in-house QC process were removed.
- Pearson correlation coefficients (with 95% confidence intervals) are given for correlations and all analyses were carried out in R (version 4.3.1), unless otherwise stated.

## 248 **RESULTS**

249 Of 1746 unique participant samples included, tranches 1&2 were designated the Discovery analysis  
250 dataset (comprising 1045 samples), and tranches 3&4 the Replication analysis dataset (701 samples).

251

### 252 **Selection of Each Normalisation Procedure**

253 To define the optimal QC pipeline, selection of appropriate standardisation procedures was needed.

254 These had been previously optimised by SomaLogic for plasma samples. We set out to test how well  
255 these procedures performed in all 1746 SF samples with 7596 SOMAmer features.

256

257 We measured the impact of technical variation in each normalisation step (see methods) on assay  
258 performance in a sequential manner, firstly by measuring effects on the mean %CV and the mean  $R^2$   
259 across all proteins for the pooled sample replicates across all plates (Figure 1A&B). Comparing each  
260 to the raw RFU data, %CVs ranged from 11.65% to 16.49% for pooled OA samples (n=22) and 10.56%  
261 to 16.63% for pooled injury samples (n=22). Mean  $R^2$  ranged from 77.70% to 88.98% and 82.59% to  
262 89.10% for pooled OA and injury samples respectively, suggesting a high level of repeatability. A  
263 decreasing trend in mean %CVs suggested that the routine normalisation steps improved measure  
264 repeatability with the exception of median normalisation (Figure 1A&B). Removing median  
265 normalisation from the standardisation procedure resulted in a mean %CV of 13.6% and a mean  $R^2$   
266 of 88.98% for pooled OA samples and %CV of 14.34% and a mean  $R^2$  of 89.1% for pooled injury  
267 samples (Figure 1A&B).

268

269 Correlation of the SomaScan assay data with nine selected analytes measured in the same SF sample  
270 by immunoassay was generally high for most analytes (Figure 1C&D). However, upon normalisation  
271 steps, we saw a similar effect, with median signal normalisation reducing the correlation with these  
272 validation measurements, particularly in the injury samples (Figure 1D). Based on these combined  
273 data, we chose to use SomaLogic's existing standardisation procedures, though omitting median  
274 normalisation from our SF standardisation pipeline (i.e. employing hybridisation normalisation, plate  
275 scaling, and plate calibration using SomaLogic's plasma calibrators).

276

### 277 **Identification and correction of confounding factors**

278 After conducting standardisation, principal component (PC) analysis identified one dominant  
279 component (PC1) that explained 48% of variation in the data (Figure 2A). This principal component  
280 was positively correlated with almost all proteins measured, with a median correlation coefficient of  
281 0.70, with the highest correlations seen with low abundance proteins (Figure 2B). We were able to

282 rule out a total protein effect, due to a low correlation (-0.039) with standard high-abundance  
283 markers such as albumin concentration. Examining the proteins that drove this signal, we found that  
284 lower protein abundance was the strongest independent predictor of correlation (Table S5) with PC1  
285 ( $p < 2.23e-308$ ), with the next most significant predictors being whether proteins were predicted not  
286 to be secreted ( $p = 3.98e-10$ ) and proteins that were identified as nuclear and not secreted ( $p = 1.64e-$   
287  $9$ ). This led to the hypothesis that PC1 was capturing an effect of intracellular proteins, perhaps  
288 reflecting cell turnover or due to the presence of microvesicles. A strong intracellular signal was  
289 confirmed by showing that PC1 was consistently reduced in spun samples when comparing paired SF  
290 samples (from the same parent SF,  $n = 18$ ) that had been split into two and either spun or left unspun  
291 immediately after joint aspiration (Figure 2C).

292

293 To quantify the contribution of intracellular proteins, we derived an Intracellular Protein Score (IPS)  
294 as the weighted sum of relative protein concentrations. For weights, we calculated a Cohen's  $d$  from  
295 the 18 paired spun and unspun samples. This score correlated very highly with PC1 (Figure 2D). We  
296 used this score in a linear regression model adjusting for the contribution of intracellular proteins,  
297 which removed the correlation between IPS and PC1 (Figure 2E), reduced the variance explained by  
298 PC1 to 16% (Figure 2F), and removed the correlation of "non-secreted nuclear protein" with PC1  
299 (Table S5).

300

301 This intracellular contribution to the SF proteome did not correlate strongly with any of our pre-  
302 defined technical confounders. However, it explained a large proportion of variation in our data,  
303 which ran the risk of swamping more subtle protein signatures or molecular endotypes, if present.  
304 We thus decided to include, in addition to the standardised data set without IPS adjustment, a co-  
305 primary dataset, applying IPS adjustment to each protein as part of our STEpUP OA Data Analysis  
306 Plan (<https://www.kennedy.ox.ac.uk/oacentre/stepup-oa/stepup-oa>).

307

308 We also found a strong 'bimodal' signal on PC2 of the data (Figure 3A&B) whereby a large number of  
309 SOMAmers ( $N = 4030$  at Benjamini-Hochberg (BH) adjusted  $p < 0.05$ ) were present at either very low  
310 or very high relative signal in a given sample. Further investigation showed that PC2 was highly  
311 correlated with the technical variable 'laboratory processing batch' ( $p < 2.2E-308$ ). Investigation of  
312 exemplar proteins displaying this behaviour showed that the bimodal signal followed sample  
313 processing order, usually (but not always) between laboratory processing batches (Figure 3C). The  
314 effect became stronger over time (Figure 3C). Re-analysing (at SomaLogic) previously laboratory  
315 processed (hyaluronidase treated) samples gave the same result (data not shown). However, when

316 new aliquots of three sequential samples which had differing bimodal status were reprocessed by  
317 the Oxford laboratory and re-analysed, all three reprocessed results had a shared bimodal status.  
318 This indicated that this was due, in some way, to our laboratory sample processing (hyaluronidase  
319 treatment) (Figure 3D). We hypothesised this might be due to sample temperature differences prior  
320 to hyaluronidase treatment, but further experiments did not corroborate this (data not shown).  
321 Neither was this thought to be due to the stability of frozen hyaluronidase enzyme as the pooled OA  
322 sample, freshly processed during each tranche with frozen stored hyaluronidase, showed little  
323 variability over time (Figure S2).

324

325 We applied a Gaussian Mixture Model to PC2 to classify samples into high or low protein status,  
326 reflecting their bimodal signal (Figure 3E), which produced visually plausible assignments on PCA and  
327 UMAP (Figure 3E&F). To attempt to reduce this undesired variance, we carried out batch correction  
328 by samples' PC2 bimodal signal status using the ComBat method<sup>[50]</sup>. This correction reduced the  
329 impact of the bimodal signal considerably (Figure 3G&H) and was adopted into our QC pipeline.

330

331 We also discovered a significant influence of plate on a number of proteins (n=1927) at BH adjusted  
332  $p < 0.05$ ). Our samples were randomised to plate, so this was unlikely to cause significant confounding  
333 in downstream analyses, but to reduce technical variation we also applied batch correction for plate  
334 by ComBat at the same time as correcting for the bimodal signal.

335

336 We assessed the impact of these adjustments described above using the immunoassay comparison  
337 data. While the IPS-adjustment reduced the dominance of PC1, it also had a negative impact on the  
338 correlation coefficients between SomaScan and prior immunoassay results of the select analytes  
339 (Figure 4). This was particularly evident in the injury group, where the correlation between the two  
340 measures for 4 out of 9 proteins (IL6, IL8, TGF $\beta$ 1, TIMP1) changed from strongly correlated to weakly  
341 or not correlated (Figure 4B). Interestingly, some of the measured cytokines (which had been  
342 selected because of their putative disease relevance) such as MCP1, IL-8 and TGF $\beta$ 1, correlated with  
343 the intracellular protein score (Table S6). Batch correction for plate and bimodal signal status (as  
344 part of our optimised standardisation) was seen to have minimal impact on immunoassay  
345 agreement (Figure 4).

346

#### 347 **Description and reduction of pre-defined technical confounding by protein filtering**

348 In addition to identifying badly performing SOMAmers and samples for filtering according to assay  
349 performance (see methods), we also identified filters based on pre-defined technical confounders

350 (Table S2).

351

352 Confounding technical factors were dealt with in different ways. Samples showed systematic biases  
353 in signal intensity plate position (Table S7, 'Plate position'), but this was also deemed unlikely to  
354 confound downstream analyses because samples were randomised across and within plates, so its  
355 effect was not adjusted for. Blood staining (Table S7, 'Visual blood staining') and sample volume  
356 were both drivers of IPS, but we felt that both could contain biological signals of relevance, so they  
357 were not adjusted during QC but were considered covariates in the downstream analyses. Sample  
358 age (Table S7, 'Sample age') could introduce technical variation, therefore significantly associated  
359 proteins were removed by filtering (Table S4). Freeze-thawing was also shown to be a potential  
360 technical confounder (Table S7, 'Sample freeze thaw cycles'). This was investigated further.

361

362 For freeze-thawing, we had repeatedly freeze-thawed (five times per sample) one aliquot each of  
363 the pooled OA and the pooled injury samples. This had an effect, particularly in the injury samples.  
364 However, the majority (77%) of proteins retained a good %CV (<20%) even after five freeze-thaws  
365 (Figure S2), suggesting that such samples remained usable. Technical variation brought about by  
366 freeze-thaw was nonetheless adjusted for by filtering out significantly associated proteins following  
367 Bonferroni correction (Table S4).

368

### 369 **Assessment of centrifugation effect on protein measurements**

370 Although most of the samples had been centrifuged prior to initial storage as per our eligibility  
371 criteria, 240 samples included in the replication analysis were unspun. In anticipation of this we  
372 assessed further the impact of centrifugation on the 18 pairs of samples that had either been spun  
373 or left unspun at time of collection. We compared the SomaScan data to identify proteins that  
374 changed upon centrifugation (Figure S3). The effect of centrifugation on the data depended on  
375 whether the data were adjusted for IPS. The unadjusted data showed that centrifugation status was  
376 a major driver of variation across the paired samples, with a significant correlation with PC1 (Figure  
377 S3A, paired t-test  $p=0.0066$ ), but after adjusting for IPS, the top PCs were no longer driven by spun  
378 status (Figure S3B, paired t-test  $p=0.2089$ ). Centrifugation impacted the concentration of a large  
379 number of individual proteins in both IPS-unadjusted ( $n=5638$ , 74%, at BH adjusted  $p<0.05$ ) and, to a  
380 lesser extent, IPS-adjusted data ( $n=3731$ , 49%, at BH adjusted  $p<0.05$ ), although the majority of  
381 proteins were significantly correlated between the paired spun and unspun samples ( $n=6402$ , 85% in  
382 unadjusted data and  $n=4558$ , 60% in IPS adjusted data, Figure S3C&D respectively).

383



384 We concluded that spun and unspun samples were comparable (in that they captured similar  
385 information), but that any analysis that included both types together would need to adjust for  
386 systematic shifts in abundance and the small numbers of uncorrelated proteins. In our discovery and  
387 replication analysis plans relating to our primary analysis, only spun samples are therefore  
388 considered, with unspun samples used for secondary sensitivity analyses.

389

#### 390 **Effect of blood staining on protein measurements**

391 A subset of samples had information on blood staining, graded by visual inspection at the time of  
392 joint aspiration, prior to centrifugation. As shown in Table S7, the presence of blood measured in this  
393 way was a significant driver of protein variation. It was also a potential biological driver as  
394 haemarthrosis is common after significant joint injury and is known to be pro-inflammatory and  
395 associated with persisting knee symptoms<sup>[51-53]</sup>. Visual blood staining could reflect presence of either  
396 intact or lysed red blood cells. The analyte haemoglobin A (HBA) correlated reasonably well with  
397 visual blood staining grade prior to adjustment for IPS (Figure S4A), and less so after adjustment for  
398 IPS (Figure S4B). The log concentration of HBA relative abundance level (without IPS adjustment)  
399 was subsequently used as a measure of blood content, as a covariate in downstream analyses.

400

#### 401 **Validation of data quality after QC**

402 Following application of filters, 1720 samples and 6290 SOMAmers (features) remained. The total  
403 numbers of samples and proteins filtered out are shown in Table S4. After filtering, median %CV of  
404 pooled OA and injury samples remained relatively unchanged at 11.25 and 12.42 respectively (Figure  
405 S5). An overview of the end-to-end data processing and quality control pipeline, from raw data to  
406 final filtered data, is shown in Figure 5.

407

408 The association of all these variables with the top 10 PCs of the standardised, bimodal signal  
409 corrected data after filtering is shown in Table S7. The strongest associations in the IPS adjusted,  
410 filtered data are shown in Figure 6. IPS adjusted non-filtered, and non-IPS adjusted data are shown  
411 in Figure S6.

412

413 Finally, we visualised the different diagnostic subgroups (OA, joint injury, inflammatory control,  
414 disease-free control) on UMAPs of the standardised, corrected and filtered data, with and without  
415 IPS adjustment (Figure 7A&B respectively). Both datasets showed clustering of knee injury and OA  
416 cases in overlapping but distinguishable areas of high-dimensional proteomic space, though the  
417 smaller groups (disease-free controls and inflammatory controls) were more evenly distributed.

418 Inflammatory controls tended to segregate with acute knee injury samples. These patterns were also  
419 reflected at the PC level (Figure 6F, Figure S6).

## 420 **DISCUSSION**

421 In this initial report from the STEpUP OA consortium, we describe a comprehensive evaluation of the  
422 overall performance of the SomaScan assay for knee SF for the first time. We address a series of data  
423 processing and analysis challenges that arise from proteomic quantification of SF using this  
424 technology. Based on our investigations, we propose an optimal standardisation procedure for SF  
425 and an assessment of the quality of the protein data using pre-defined approaches. Our aim was to  
426 justify the best approach to minimize technical variation while maintaining biological variation in the  
427 data, and thus to develop a pipeline that could be applied to downstream analyses within STEpUP  
428 OA and in subsequent proteomic analyses of SF by others.

429

430 We identified a number of technical confounders, which all affected, to a great or lesser extent, the  
431 data structure. These included factors that related to the SF sample e.g. its age, number of freeze-  
432 thaws, as well as potential confounders that could arise during the sample processing e.g. plate and  
433 date of processing. Because we had randomised the samples to and within plate, we were also able  
434 to identify plate position and laboratory processing batch as additional confounding factors. Each of  
435 these was controlled for by adjustment or filtering (either sample or SOMAmer). Whilst our  
436 intention is to perform the primary analysis of STEpUP OA only in spun SF sample data, we also  
437 included a number of unspun samples to test the generalizability of these to the larger dataset. In  
438 doing so we calculated that more stringent filtering would be needed when studying unspun, or  
439 mixed spun/unspun SF collections. We chose not to correct for blood staining (or HBA) or sample  
440 volume as we felt that these could reflect important biological variation.

441

442 The finding that a large proportion of variance in the data was driven by intracellular protein was  
443 unexpected. This could have arisen as a result of technical confounding following contamination of  
444 the SF samples by cells e.g. by a failure to remove cells fully by centrifugation, or by cell lysis of those  
445 cells at the time of aspiration e.g. by delay in spinning sample down. It could also be a true reflection  
446 of cellular turnover within the joint as part of the disease process e.g. of infiltrating immune cells or  
447 native connective tissues. It could also reflect protein carried within microvesicles that are known to  
448 be increased in joint disease and which drive biology within and between joint tissues<sup>[54, 55]</sup>. These  
449 possibilities are currently being explored. There is a concern that if the IPS reflects true biology, then  
450 adjusting for it may compress true signals within the data. This is consistent with the reduction in  
451 correlation with immunoassay seen in the injury samples. On the other hand, subtle structures  
452 within the data that reflect true molecular endotypes might be masked without removal of this

453 signal. For this reason, the relevance of the IPS to clinical parameters and endotype clusters will be  
454 addressed alongside one another in the primary analysis of STEpUP OA.

455

456 Our data demonstrate that, when properly processed, SomaScan produces repeatable and accurate  
457 quantification of SF proteins which can (as a quality check, not a diagnostic one) broadly separate  
458 different clinical groupings. Its quantitation quality is at least comparable/superior to other 'non-  
459 standard' matrices e.g. urine/cerebrospinal fluid previously studied on this platform<sup>[56-58]</sup>. Our  
460 repeatability measures in pooled synovial fluid samples, with median %CVs of 11.25% for OA sample  
461 replicates and 12.42% for injury sample replicates, were higher than have been reported for plasma  
462 and serum samples measured using SomaScan technology, where %CVs of 5% or less are  
463 observed<sup>[59]</sup>. However, our correlation with immunoassays (median coefficient 0.81 for OA and 0.92  
464 injury respectively) are as good or better than are typically observed in blood samples<sup>[60, 61]</sup> . Our  
465 data quality metrics are comparable with those based on mass spectrometry or other quantitative  
466 immunoassays, where %CVs of 10% or greater are recorded<sup>[62, 63]</sup>. Compared with these  
467 technologies our data has higher dynamic range and sensitivity (as the technology can assay proteins  
468 that are at very high as well as low abundance), noting only proteins on a pre-defined (though very  
469 large) protein list are included.

470

471 There are several limitations of this large study. The consortium collection was highly  
472 heterogeneous, gathered from seventeen different studies, varying in disease severity and  
473 phenotype, across several decades and from a number of countries without a unified pre-specified  
474 sample processing protocol. This made distinguishing technical and biological variation difficult (as  
475 was the case for the intracellular protein score, which could reflect either variation in joint biology or  
476 variation in sampling handling). A further limitation is that we analysed only a single matrix (SF)  
477 although this likely reflects activity in multiple joint tissues. The lack of paired  
478 cartilage/synovium/bone in STEpUP OA prevents us performing a direct integrated analysis using  
479 RNAseq, for example, although it may be possible to extrapolate this from other existing datasets.  
480 Other proteome-wide technologies (such as LC-MS/MS, OLINK<sup>[64]</sup>) could provide further validation  
481 on protein patterns within OA SF. Paired plasma is also available for many individuals in STEpUP OA  
482 but is yet to be analysed. From previous experience we would predict that this matrix would show  
483 low concordance with SF<sup>[41, 65]</sup>. Others have used SomaScan to explore the plasma proteome in OA,  
484 identifying diagnostic and prognostic biomarkers, though this study did not study paired SF  
485 samples<sup>[34]</sup> .

486

487 In summary, we present an evidence-based methodology pipeline for large scale proteomic analysis  
488 on the SomaScan platform of SF, which has the potential to be a critical matrix for discovery science  
489 and clinical translation in OA. Our next step, the primary analysis of this dataset, seeks to answer  
490 definitively whether there are distinct discernible molecular endotypes in this common, yet poorly  
491 understood disease.  
492

493 **Funding Statement:**

494 The study was supported by Kennedy Trust for Rheumatology Research (grant number: 171806),  
495 Versus Arthritis (grant number: 22473), Centre for OA Pathogenesis Versus Arthritis (grant numbers:  
496 21621, 20205), Galapagos, Biosplice, Novartis, Fidia, UCB, Pfizer (non consortium member) and  
497 Somalogic (in kind contributions).

498

499 **Competing Interest Statement:**

500 YD, TAP, PH, SL, AS, NKA, DF, BM, AMV, SK, VB, JMA and VK declare no conflicts of interest. FW has  
501 received consultancy fees from Pfizer, and has a leadership role at the Medical Research Council  
502 (panel member) and Osteoarthritis and Cartilage (Associate Editor). LSL has received consultancy  
503 fees from Arthro Therapeutics AB, and is an advisory board member of AstraZeneca (non consortium  
504 member). LJD has received consultancy fees from Nightingale Health PLC. TLV has no conflicts to  
505 declare with the exception of grant income for STEpUP OA from industry partners (see above). RAM  
506 is a shareholder of AstraZeneca. SB and JM are employees and shareholders of Novartis (consortium  
507 members). MK has received support for attending the Gordon Research Conference, OARSI meeting,  
508 International Cartilage Repair Society, Munster University, is a board member of the Dutch Arthritis  
509 Society (Chair of Visitation Board), and has a leadership role at Osteoarthritis Research Society  
510 International (Board of Directors Member). DAW has received consultancy fees from  
511 GlaxoSmithKline plc, AKL Research & Development Limited, Pfizer Ltd, Eli Lilly and Company, Contura  
512 International, and AbbVie Inc, has received honoraria for educational purposes from Pfizer Ltd and  
513 AbbVie Inc, is a board member of UKRI (Director) and Versus Arthritis Advanced Pain Discovery  
514 Platform.

515

516 **Author contributions:**

517 Conception and Design: TLV, FEW, LJD, PH, RAM, DP, SL, SB, LSL, AS, CTA, DF, BDMT, MK, TJW, DAW,  
518 AMV. Analysis and interpretation of data: YD, TAP, LJD, FEW, TLV, PH, RAM, JM, SB, BDMT, LB.  
519 Drafting Article: TLV, TAP, YD, LJD, FEW. Critical revision of article: all authors. Final Approval: all  
520 authors.

521

522 **Acknowledgements:**

523 We would like to express our gratitude and thanks to all cohorts and their participants who  
524 contributed samples to STEpUP OA. We are grateful for the support from Floris Lafeber and Simon  
525 Mastbergen (Utrecht Medical Centre). This work was also supported by the NIHR Oxford Biomedical  
526 Research Centre (BRC) and the NIHR Nottingham BRC. The views expressed are those of the authors  
527 and not necessarily those of the NHS, the NIHR or the Department of Health. Tissue samples and/or  
528 data obtained from the Oxford Musculoskeletal Biobank were collected with informed donor  
529 consent in full compliance with national and institutional ethical requirements, the UK Human Tissue  
530 Act, and the Declaration of Helsinki (HTA Licence 12217 and Oxford REC C 09/H0606/11). We thank  
531 the Oxford Knee Surgery Team including Andrew Price, William Jackson and Nicholas Bottomley and  
532 our centre tissue coordinators Louise Hill and Katherine Groves who coordinated this study. We  
533 thank Charlotte Kerr for her administrative support of the consortium at large.

534

535 The STEpUP OA Consortium author block includes: University of Nottingham: Ana M. Valdes, David  
536 A. Walsh, Michael Doherty, Vasileios Georgopoulos; Lund University: Staffan Larsson, L. Stefan

537 Lohmander, André Struglics; University of Cambridge: Brian D.M. Tom, Laura Bondi; University of  
538 Toronto: Mohit Kapoor, Rajiv Gandhi, Anthony Perruccio, Y. Raja Rampersaud, Kim Perry; University  
539 of Manchester: Tim Hardingham, David Felson; University of Oxford: Tonia L. Vincent, Thomas A.  
540 Perry, Luke Jostins-Dean, Yun Deng, Vicky Batchelor, Jennifer Mackay-Alderson, Gretchen Brewer,  
541 Rose M. Maciewicz, Brian Marsden, Nigel K. Arden, Philippa Hulley, Andrew Price, Stefan Kluzek,  
542 Megan Goff, Vinod Kumar, James Tey; Imperial College London: Fiona E. Watt, Andrew Williams,  
543 Artemis Papadaki; University College Maastricht: Tim J. Welting, Pieter Emans, Tim Boymans,  
544 Liesbeth Jutten, Marjolein Caron, Guus van den Akker; University of Western Ontario: C. Thomas  
545 Appleton, Trevor B. Birmingham, J. Daniel Klapak; Biosplice: Sarah Kennedy, Jeymi Tambiah; Fidia:  
546 Devis Galesso, Nicola NK; SomaLogic: Joe Gogain, Darryl Perry, Anna Mitchel, Ela Zepko; Novartis:  
547 Sophie Brachat, Joanna Mitchelmore, Juerg Gasser, Lori Jennings; UCB: Waqar Ali.

548

549 TLV directs the Centre for OA pathogenesis (grant numbers 21612 and 20205) and has additional  
550 grant support from Versus Arthritis, the European Research Council, the Medical Research Council  
551 and FOREUM. LJD is supported by a Wellcome trust fellowship grant 208750/Z/17/Z and Kennedy  
552 Trust for Rheumatology Research for the present manuscript. LJD is also supported by grants from  
553 the MRC and the Helmsley Charitable Trust. FEW is supported by a UKRI Future Leaders Fellowship  
554 (MRC number: MR/S016538/1 and MR/S016538/2). FW, NKA and SK are members of the Centre for  
555 Sport, Exercise and Osteoarthritis Research Versus Arthritis (grant number 21595). MK is supported  
556 by grants from CIHR, NSERC, The Arthritis Society Canada, Krembil Foundation, CFI, Canada Research  
557 Chairs program, and has received support from the University Health Network Foundation, Toronto  
558 for the present manuscript. TJW is supported by grants from NWO-TTW Perspectief (#P15-23),  
559 Stichting de Weijerhorst and ReumaNederland (LLP14) for the present manuscript, and is a  
560 shareholder of Chondropeptix BV. BDMT is supported through the United Kingdom Medical  
561 Research Council programme (grant MC UU 00002/2). For the purpose of open access, the authors  
562 have applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript  
563 version arising. LB is supported by grants from Kennedy Trust for Rheumatology Research (grant  
564 number 171806) and UK Medical Research Council (grant MC UU 00002/2). DAW is supported by  
565 grants from Pfizer Ltd, UCB Pharma, Orion Corporation, GlaxoSmithKline Research and  
566 Development, and Eli Lilly and Company, Versus Arthritis, UKRI, Nuffield Foundation.

567

#### 568 **Data Access:**

569 In accordance with the STEpUP OA Consortium Agreement and the Data Access and Publication  
570 Group,  
571 protein and clinical data will be available for bone fide research relating to osteoarthritis through an  
572 application to the STEpUP OA Data Access and Publication group once the discovery and replication  
573 analyses are in press and if it does not infringe patent position. This may be subject to an access fee  
574 (to be confirmed).

575

#### 576 **Ethical Approval:**

577 The ethical approval reference numbers for individual participating cohorts are provided in Table S1.  
578 In addition, a University CUREC approval was granted for the study (details in Methods).

579

#### 580 **Supplementary Data:**

581 Code availability [https://github.com/dengyun-git/STEpUp\\_QC\\_Paper](https://github.com/dengyun-git/STEpUp_QC_Paper). For access to primary data

582 used in this analysis, see Data Access section.

583

584 **Patient and Public Involvement Statement:**

585 People with lived experience of osteoarthritis have been involved in the design of this project. A  
586 patient research panel was involved in discussing and inputting on the STEpUP OA project in  
587 February 2020 (invited to the Centre for Osteoarthritis Pathogenesis Versus Arthritis in Oxford, as  
588 part of its involvement activities). Aspects relevant to the development of the project were further  
589 discussed with the panel in July 2022. The working groups for the consortium include one focused  
590 on patient involvement and engagement. A lay summary is included in the appendix of our publicly  
591 available analysis plan. A short video about the project was produced and is available on our  
592 website: <https://www.kennedy.ox.ac.uk/oacentre/stepup-oa>. In addition, the various constituent  
593 cohorts contributing to STEpUP OA also typically have lay or patient members on their steering  
594 committees.



595 **Figures/illustrations:**

596

597 **Figure 1.** Assessment of the effects of each standardisation step on (A) mean %CV and (B) mean  
598  $R^2$  across all proteins for pooled sample replicates, stratified by OA and acute knee injury  
599 respectively. Assessment of Pearson correlation coefficients between protein expression in samples  
600 measured by the SOMAscan platform and by prior immunoassay for nine select proteins across  
601 normalisation steps for (C) OA and (D) acute knee injury. The normalisation steps included  
602 hybridisation normalisation (HN), plate scaling using plasma calibrators (PS), median signal  
603 normalisation (MN) and calibration using plasma calibrators (PC). Correlation between the RFUs  
604 (SOMAscan) and absolute concentrations for the nine proteins across the two methods are shown.  
605 RFUs, relative fluorescence units; %CV, % coefficient of variation; osteoarthritis, OA; Activin A, Inhibin  
606 beta A chain; FGF2, Fibroblast growth factor 2; IL6, Interleukin-6; IL8, Interleukin-8; MCP1, C-C motif  
607 chemokine 2; MMP3, Stromelysin-1; TGF $\beta$ 1, Transforming growth factor beta-1; TIMP1, Tissue  
608 inhibitor of metalloproteinase 1; TSG6, Tumor necrosis factor-inducible gene 6.

609

610 **Figure 2.** (A) Variation explained (%) by the top 10 PCs derived from the standardised log abundance  
611 proteomic data. (B) Correlation between PC1 and protein abundance, with two high-abundance  
612 proteins (albumin, a soluble serum protein, and LDH, an intracellular protein) marked. Protein  
613 abundance is calculated as the standardized RFU for each protein adjusted by the protein's dilution  
614 factor used in the SomaScan assay (the "dilution bin"). (C) Comparison of variation explained (%) by  
615 PC1 between 18 pairs of SF samples that were centrifuged (spun) or not (unspun) after aspiration  
616 and prior to freezing, with paired samples from the same participant joined by separate lines. Red  
617 lines show samples that had an increased PC1 prior to spinning, and the green line where it was  
618 decreased. Correlation between PC1 and intracellular protein score (D) before and (E) after IPS  
619 adjustment. (F) Variation explained by the top 10 PCs derived from the batch corrected and IPS  
620 adjusted log abundance proteomic data. In all cases, correlation is measured using the Pearson  
621 correlation coefficient. IPS, Intracellular Protein Score; PC, principal component; LDH, Lactate  
622 dehydrogenase.

623

624 **Figure 3.** (A) Distribution of the second principal component (PC2) derived from the standardised log  
625 abundance data, showing a bimodal distribution. (B) UMAP visualisation of two reduced dimensions  
626 (D1 and D2) of the top PCs of the standardised log abundance data. (C) Example of a strongly  
627 bimodal protein measurement, TSG101, RFU (y-axis) against Oxford laboratory processing order (x-  
628 axis) and coloured by laboratory processing batch (with only points within the same processing batch  
629 connected by lines). Note that the 'flipping' between high and low signal status occurred primarily  
630 when processing batch changed, and only rarely within processing batch. This effect was particularly  
631 strong among sample batches that were processed later in processing order. (D) The same example  
632 protein measurement for three independent SF samples before (original) and after they were re-  
633 processed and re-assayed, showing that bimodal status changed after laboratory re-processing. (E)  
634 Distribution of PC2 derived from standardised log abundance data, showing the two probability  
635 density functions of the Gaussian Mixture Model used to classify samples into the two bimodal signal  
636 status groups. (F) UMAP visualisation of two reduced dimensions (D1 and D2) of the top PCs of the  
637 standardised log abundance data, colored by the inferred bimodal signal status. (G) Histogram of PC2  
638 of the batch corrected log abundance data, with the now near-identical distributions of the two  
639 bimodal signal status groups shown as colored lines, (H) UMAP visualisation on two reduced

640 dimensions (D1 and D2) of the top PCs of the batch corrected log abundance data, colored by the  
641 inferred bimodal signal status. RFUs, relative fluorescence units; PC, Principal Component; TSG101,  
642 Tumor susceptibility gene 101 protein; UMAP, Uniform Manifold Approximation and Projection.

643

644 **Figure 4.** Correlation between SOMAscan relative frequency abundance (RFU) and abundance  
645 measured using orthogonal immunoassays for 9 selected proteins at different stages of  
646 SOMAscan data processing, for (A) osteoarthritis and (B) acute knee injury samples. Correlation was  
647 measured using the Pearson correlation coefficient. Raw data refers to the raw RFUs without any  
648 processing, optimised standardisation was the data standardised using our selected optimal  
649 normalization steps (Figure 1), processed without IPS adjustment refers to data that has been batch  
650 corrected for bimodal signal status and plate but not IPS adjusted, and processed with IPS  
651 adjustment refers to samples that have undergone both batch correction and IPS adjustment. IPS,  
652 Intracellular Protein Score; Protein name abbreviations as in Figure 1.

653

654 **Figure 5.** Overview of the final data processing and quality control pipeline for synovial fluid  
655 SOMAscan data used by the STEpUP OA consortium, broken down into three stages: standardisation  
656 (yellow box), technical confounder correction (blue box) and filtering (green box). More details on  
657 filtering thresholds, and the number removed by each filter, can be found in Supplementary Table S4.

658

659 **Figure 6.** Visualisation of selected pre-defined confounders against select principal components of  
660 the batch corrected, filtered, IPS adjusted data. (A) The average value of PC9 (most strongly  
661 associated with plate position) by sample well position, (B-F) visualisation of the two PCs most  
662 strongly associated with each confounder, coloured by confounder value. Pre-defined confounders  
663 shown are (B) blood staining grade of sample after aspiration assessed by visual inspection, (C)  
664 volume of sample taken during aspiration, (D) age of the sample in years, measured from aspiration  
665 to sample processing at Oxford, (E) the number of times the sample was thawed and re-  
666 frozen before sample processing at Oxford, (F) the disease group of the sample (osteoarthritis [OA],  
667 acute knee injury [Injury], healthy control, inflammatory arthritis control).

668

669 **Figure 7.** UMAP visualisation of two reduced dimensions (D1 and D2) of the top PCs of the log  
670 abundance data with (A) and without (B) IPS adjustment followed by filtering, coloured by  
671 disease group. These groups were osteoarthritis (OA, acute knee injury (injury), healthy controls,  
672 inflammatory arthritis controls.

673 UMAP, Uniform Manifold Approximation and Projection.

674

675 **Figure S1.** Consortium structure, as working groups. Distinct working groups oversaw key activities  
676 according to pre-defined Terms of Reference (available on request). TV, Tonia Vincent; FW, Fiona  
677 Watt; AV, Ana Valdes; LJD, Luke Jostins-Dean; RM, Rose Maciewicz.

678

679 **Figure S2.** Assessment of assay repeatability using pooled samples of synovial fluid from participants  
680 with (A) knee OA and (B) acute knee injury, measured by the coefficient of variation (%CV). These  
681 include the repeatability of the standard processed pooled samples included on every plate ('Sample  
682 Repeats'), pooled samples which had been repeatedly freeze-thawed ('Freeze Thaw') prior to  
683 processing and an OA pool aliquot that had been freshly enzyme digested with stored hyaluronidase

684 during each of the 2nd, 3rd and 4th tranches of sample processing (done for the OA pool only)  
685 ('Reprocessed'). Dotted vertical lines show the maximum %CV for 80% of proteins for each group.

686

687 **Figure S3.** Top 2 principal components of (A) non-IPS adjusted and (B) IPS adjusted log RFU of the 18  
688 pairs of centrifuged (spun) and non-centrifuged (unspun) SF samples. Samples are coloured by spin  
689 status and paired samples are linked by lines. Measures of differential abundance (Cohen's d) and  
690 Pearson correlation coefficient ( $\rho$ ) between spun and unspun samples for (C) non-IPS adjusted and  
691 (D) IPS adjusted log RFU. Samples are coloured depending on their significance (Benjamini-Hochberg  
692 adjusted  $p < 0.05$ ) on the two measures: Different Means corresponds to a significant difference in  
693 means in a paired t-test and Correlated corresponds to a significant correlation in a Pearson  
694 correlation test. IPS, intracellular protein score; PC, principal component; SF, synovial fluid; RFU,  
695 relative fluorescence unit.

696

697 **Figure S4.** Boxplots showing the correlation between visual blood staining grade of SF at the time of  
698 sample collection and the blood analyte, HBA, in non-IPS adjusted data in (A) all samples, (C) OA  
699 samples and (E) acute knee injury samples, and in IPS adjusted data in (B) all samples, (D) OA  
700 samples and (F) acute knee injury samples. Spearman correlation coefficients measuring rank-based  
701 correlation considering visual blood staining as an ordinal variable are shown. 443 OA samples had  
702 blood staining grade 1 (no blood detected, 75% among the 588 total samples with blood staining  
703 records). HBA, haemoglobin A; IPS, intracellular protein score; SF, synovial fluid.

704

705 **Figure S5.** (A) Assessment of assay repeatability after optimised quality control procedures measured  
706 using the cumulative distribution of the coefficient of variation (%CV) on pooled OA samples (OA  
707 Sample Repeats) and pooled acute knee Injury samples (Injury Sample Repeats) separately. 80% of  
708 proteins had a %CV less than 16.85% and 17.57% in the OA and acute knee injury pools (blue and red  
709 dotted lines respectively). (B) The proportion of variation that was estimated to be non-technical,  
710 measured by  $R^2$  for OA and acute knee injury sample repeats separately. 80% of proteins had  $R^2$   
711 values greater than 88.27% and 84.33% in the OA and knee injury pools (blue and red dotted lines  
712 respectively).

713

714 **Figure S6.** Visualisation of pre-defined technical confounders by select principal components of the  
715 (A) non-filtered IPS adjusted (B) filtered non-IPS adjusted and (C) non-filtered non-IPS adjusted data.  
716 Visualisation of the two PCs most strongly associated with each confounder (colours correspond to  
717 confounder value). Confounders include plate position (mean of PC8), blood staining grade of sample  
718 (which was performed immediately after aspiration from the joint by visual inspection), volume of  
719 sample taken during aspiration, age of the sample in years from aspiration to processing, the number  
720 of times the sample had been thawed and re-frozen, the disease group of the sample (knee  
721 OA, acute knee injury, healthy control, inflammatory arthritis control). The association between  
722 each PC and confounder is shown in the Table S7.

723

724 **Figure S7.** Pairwise scatter plots (off-diagonal) and histograms (diagonal) of the top five principal  
725 components of standardised log abundance, (A) before and (B) after batch correction for plate and  
726 bimodal signal status, coloured by bimodal signal status. Batch correction effectively removed the  
727 effect of bimodal signal status on the top PCs.

728 References:

729

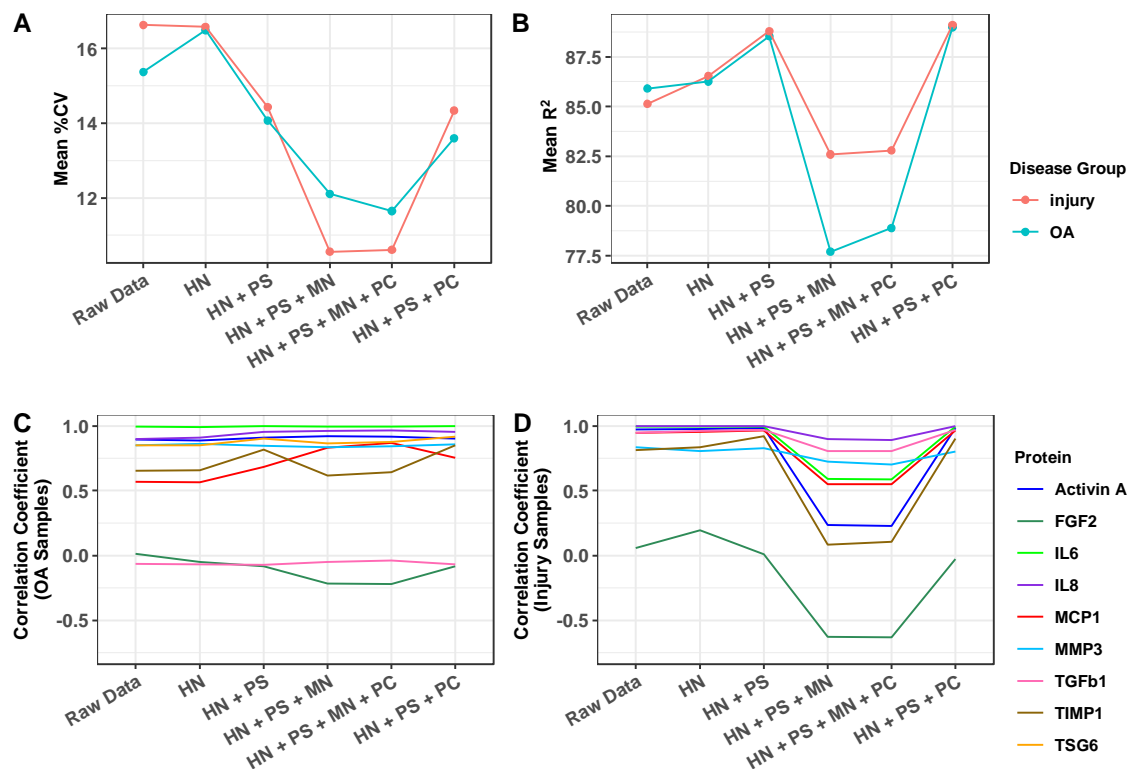
- 730 1. Hunter, D.J., L. March, and M. Chew, *Osteoarthritis in 2020 and beyond: a Lancet*  
731 *Commission*. Lancet (London, England), 2020. **396**(10264): p. 1711-1712.
- 732 2. Safiri, S., et al., *Global, regional and national burden of osteoarthritis 1990-2017: a*  
733 *systematic analysis of the Global Burden of Disease Study 2017*. Annals of the rheumatic  
734 diseases, 2020. **79**(6): p. 819-828.
- 735 3. *OARSI White Paper- OA as a Serious Disease*. 2016 [cited 07-07-23]; Available from:  
736 <https://oarsi.org/education/oarsi-resources/oarsi-white-paper-oa-serious-disease>.
- 737 4. Hawker, G.A., *Osteoarthritis is a serious disease*. Clinical and experimental rheumatology,  
738 2019. **37 Suppl 120**(5): p. 3-6.
- 739 5. Lane, N.E., et al., *OARSI-FDA initiative: defining the disease state of osteoarthritis*.  
740 Osteoarthritis and cartilage, 2011. **19**(5): p. 478-82.
- 741 6. (NICE), N.I.f.H.a.C.E. *Osteoarthritis: care and management*. 2020 [cited 07-07-23]; Available  
742 from: <https://www.nice.org.uk/guidance/cg177/chapter/1-Recommendations#diagnosis-2>.
- 743 7. Zhang, W., et al., *EULAR evidence-based recommendations for the diagnosis of knee*  
744 *osteoarthritis*. Annals of the rheumatic diseases, 2010. **69**(3): p. 483-9.
- 745 8. Vincent, T.L., *Of mice and men: converging on a common molecular understanding of*  
746 *osteoarthritis*. The Lancet Rheumatology, 2020. **2**(10): p. e633-e645.
- 747 9. Mobasheri, A., et al., *Molecular taxonomy of osteoarthritis for patient stratification, disease*  
748 *management and drug development: biochemical markers associated with emerging clinical*  
749 *phenotypes and molecular endotypes*. Curr Opin Rheumatol, 2019. **31**(1): p. 80-89.
- 750 10. Karsdal, M.A., et al., *Disease-modifying treatments for osteoarthritis (DMOADs) of the knee*  
751 *and hip: lessons learned from failures and opportunities for the future*. Osteoarthritis  
752 Cartilage, 2016. **24**(12): p. 2013-2021.
- 753 11. Hochberg, M.C., et al., *Effect of Intra-Articular Sprifermin vs Placebo on Femorotibial Joint*  
754 *Cartilage Thickness in Patients With Osteoarthritis: The FORWARD Randomized Clinical Trial*.  
755 JAMA, 2019. **322**(14): p. 1360-1370.
- 756 12. Eckstein, F., et al., *Long-term structural and symptomatic effects of intra-articular sprifermin*  
757 *in patients with knee osteoarthritis: 5-year results from the FORWARD study*. Annals of the  
758 Rheumatic Diseases, 2021. **80**(8): p. 1062-1069.
- 759 13. Kloppenburg, M. and W.Y. Kwok, *Hand osteoarthritis--a heterogeneous disorder*. Nat Rev  
760 Rheumatol, 2011. **8**(1): p. 22-31.
- 761 14. Deveza, L.A. and R.F. Loeser, *Is osteoarthritis one disease or a collection of many?*  
762 *Rheumatology*, 2018. **57**: p. 34-42.
- 763 15. Driban, J.B., et al., *Is osteoarthritis a heterogeneous disease that can be stratified into*  
764 *subsets?* Clinical Rheumatology, 2010. **29**(2): p. 123-131.
- 765 16. Dieppe, P.A. and L.S. Lohmander, *Pathogenesis and management of pain in osteoarthritis*.  
766 Lancet, 2005. **365**(9463): p. 965-973.
- 767 17. Liu-Bryan, R. and R. Terkeltaub, *Emerging regulators of the inflammatory process in*  
768 *osteoarthritis*. Nature Reviews Rheumatology, 2015. **11**(1): p. 35-44.
- 769 18. Vincent, T.L., *Mechanoflammation in osteoarthritis pathogenesis*. Seminars in Arthritis and  
770 Rheumatism, 2019. **49**: p. S36-S38.
- 771 19. Vincent, T.L., *Of Mice and Men; converging on a common molecular understanding of*  
772 *Osteoarthritis*. Lancet Rheumatology, 2020. **2**(10): p. E633-E645, .
- 773 20. Luo, Y.Y., et al., *A low cartilage formation and repair endotype predicts radiographic*  
774 *progression of symptomatic knee osteoarthritis*. Journal of Orthopaedics and Traumatology,  
775 2021. **22**(1).
- 776 21. McCulloch, K., G.J. Litherland, and T.S. Rai, *Cellular senescence in osteoarthritis pathology*.  
777 Aging cell, 2017. **16**(2): p. 210-218.

- 778 22. Elahi, S., et al., *The association between varus-valgus alignment and patellofemoral*  
779 *osteoarthritis*. *Arthritis and Rheumatism*, 2000. **43**(8): p. 1874-1880.
- 780 23. Felson, D.T., et al., *The Prevalence of Knee Osteoarthritis in the Elderly - the Framingham*  
781 *Osteoarthritis Study*. *Arthritis and Rheumatism*, 1987. **30**(8): p. 914-918.
- 782 24. McAlindon, T.E., et al., *Level of physical activity and the risk of radiographic and*  
783 *symptomatic knee osteoarthritis in the elderly: The Framingham Study*. *American Journal of*  
784 *Medicine*, 1999. **106**(2): p. 151-157.
- 785 25. Sharma, L., et al., *The mechanism of the effect of obesity in knee osteoarthritis - The*  
786 *mediating role of malalignment*. *Arthritis and Rheumatism*, 2000. **43**(3): p. 568-575.
- 787 26. Doherty, M., *Risk factors for progression of knee osteoarthritis*. *Lancet*, 2001. **358**(9284): p.  
788 775-776.
- 789 27. Felson, D.T., *Identifying different osteoarthritis phenotypes through epidemiology*.  
790 *Osteoarthritis and Cartilage*, 2010. **18**(5): p. 601-604.
- 791 28. Mobasher, A., et al., *Recent advances in understanding the phenotypes of osteoarthritis*.  
792 *F1000Res*, 2019. **8**.
- 793 29. Deveza, L.A., et al., *Knee osteoarthritis phenotypes and their relevance for outcomes: a*  
794 *systematic review*. *Osteoarthritis Cartilage*, 2017. **25**(12): p. 1926-1941.
- 795 30. Nelson, A.E., et al., *Biclustering reveals potential knee OA phenotypes in exploratory*  
796 *analyses: Data from the Osteoarthritis Initiative*. *PLoS One*, 2022. **17**(5): p. e0266964.
- 797 31. Dell'Isola, A., et al., *Identification of clinical phenotypes in knee osteoarthritis: a systematic*  
798 *review of the literature*. *BMC musculoskeletal disorders*, 2016. **17**(1): p. 425-425.
- 799 32. Angelini, F., et al., *Osteoarthritis endotype discovery via clustering of biochemical marker*  
800 *data*. *Ann Rheum Dis*, 2022. **81**(5): p. 666-675.
- 801 33. Werdyani, S., et al., *Endotypes of primary osteoarthritis identified by plasma metabolomics*  
802 *analysis*. *Rheumatology*, 2021. **60**(6): p. 2735-2744.
- 803 34. Styrkarsdottir, U., et al., *The CRTAC1 Protein in Plasma Is Associated With Osteoarthritis and*  
804 *Predicts Progression to Joint Replacement: A Large-Scale Proteomics Scan in Iceland*. *Arthritis*  
805 *Rheumatol*, 2021. **73**(11): p. 2025-2034.
- 806 35. Bay-Jensen, A.C., et al., *The Need for Predictive, Prognostic, Objective and Complementary*  
807 *Blood-Based Biomarkers in Osteoarthritis (OA)*. *Ebiomedicine*, 2016. **7**: p. 4-6.
- 808 36. Watt, F.E., et al., *The molecular profile of synovial fluid changes upon joint distraction and is*  
809 *associated with clinical response in knee osteoarthritis*. *Osteoarthritis and Cartilage*, 2020.  
810 **28**(3): p. 324-333.
- 811 37. Lohmander, L.S., et al., *Changes in joint cartilage aggrecan after knee injury and in*  
812 *osteoarthritis*. *Arthritis and Rheumatism*, 1999. **42**(3): p. 534-544.
- 813 38. Struglics, A., et al., *Human osteoarthritis synovial fluid and joint cartilage contain both*  
814 *aggrecanase- and matrix metalloproteinase-generated aggrecan fragments*. *Osteoarthritis*  
815 *and Cartilage*, 2006. **14**(2): p. 101-113.
- 816 39. Catterall, J.B., et al., *Changes in serum and synovial fluid biomarkers after acute injury*  
817 *(NCT00332254)*. *Arthritis Research & Therapy*, 2010. **12**(6).
- 818 40. Timur, U.T., et al., *Identification of tissue-dependent proteins in knee OA synovial fluid*.  
819 *Osteoarthritis Cartilage*, 2021. **29**(1): p. 124-133.
- 820 41. Watt, F.E., et al., *Acute Molecular Changes in Synovial Fluid Following Human Knee Injury:*  
821 *Association With Early Clinical Outcomes*. *Arthritis Rheumatol*, 2016. **68**(9): p. 2129-40.
- 822 42. Jayadev, C., et al., *Synovial fluid fingerprinting in end-stage knee osteoarthritis: a novel*  
823 *biomarker concept*. *Bone Joint Res*, 2020. **9**(9): p. 623-632.
- 824 43. Gold, L., et al., *Advances in human proteomics at high scale with the SOMAscan proteomics*  
825 *platform*. *N Biotechnol*, 2012. **29**(5): p. 543-9.
- 826 44. Brouwers, H., et al., *Hyaluronidase treatment of synovial fluid is required for accurate*  
827 *detection of inflammatory cells and soluble mediators*. *Arthritis Res Ther*, 2022. **24**(1): p. 18.

- 828 45. Gold, L., et al., *Aptamer-based multiplexed proteomic technology for biomarker discovery*.  
829 PLoS One, 2010. **5**(12): p. e15004.
- 830 46. Schneider, D.J., et al., *Chapter 8 - SOMAmer reagents and the SomaScan platform: Chemically modified aptamers and their applications in therapeutics, diagnostics, and proteomics*, in *RNA Therapeutics*, P.H. Giangrande, V. de Franciscis, and J.J. Rossi, Editors. 2022, Academic Press. p. 171-260.
- 833 47. Harris, P.A., et al., *The REDCap consortium: Building an international community of software platform partners*. J Biomed Inform, 2019. **95**: p. 103208.
- 834 48. Georgopoulos, V., et al., *The patient acceptable symptom state for knee pain - a systematic literature review and meta-analysis*. Osteoarthritis and Cartilage, 2021. **29**: p. S52.
- 835 49. Candia, J., et al., *Assessment of Variability in the SOMAScan Assay*. Scientific Reports, 2017. **7**.
- 836 50. *ComBat: Adjust for batch effects using an empirical Bayes framework*. 2022 [cited 07-07-23]; Available from: <https://rdrr.io/bioc/sva/man/ComBat.html>.
- 837 51. Sward, P., et al., *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 Cartilage, 2012. **20**(11): p. 1302-8.
- 838 52. Sward, P., et al., *Soft tissue knee injury with concomitant osteochondral fracture is associated with higher degree of acute joint inflammation*. Am J Sports Med, 2014. **42**(5): p. 1096-102.
- 839 53. Garriga, C., et al., *Clinical and molecular associations with outcomes at 2 years after acute knee injury: a longitudinal study in the Knee Injury Cohort at the Kennedy (KICK)*. Lancet Rheumatology, 2021. **3**(9): p. E648-E658.
- 840 54. Asghar, S., et al., *Exosomes in intercellular communication and implications for osteoarthritis*. Rheumatology (Oxford), 2020. **59**(1): p. 57-68.
- 841 55. Withrow, J., et al., *Extracellular vesicles in the pathogenesis of rheumatoid arthritis and osteoarthritis*. Arthritis Res Ther, 2016. **18**(1): p. 286.
- 842 56. Helfand, B.T., et al., *A Novel Proteomics Approach to Identify Serum and Urinary Biomarkers and Pathways that Associate with Lower Urinary Tract Symptoms in Men and Women: Pilot Results of the Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN) Study*. Urology, 2019. **129**: p. 35-42.
- 843 57. Russell, T.M., et al., *Potential of High-Affinity, Slow Off-Rate Modified Aptamer Reagents for Mycobacterium tuberculosis Proteins as Tools for Infection Models and Diagnostic Applications*. J Clin Microbiol, 2017. **55**(10): p. 3072-3088.
- 844 58. Dammer, E.B., et al., *Multi-platform proteomic analysis of Alzheimer's disease cerebrospinal fluid and plasma reveals network biomarkers associated with proteostasis and the matrisome*. Alzheimers Res Ther, 2022. **14**(1): p. 174.
- 845 59. Candia, J., et al., *Assessment of variability in the plasma 7k SomaScan proteomics assay*. Sci Rep, 2022. **12**(1): p. 17147.
- 846 60. Liu, R.X., et al., *Comparison of proteomic methods in evaluating biomarker-AKI associations in cardiac surgery patients*. Transl Res, 2021. **238**: p. 49-62.
- 847 61. Raffield, L.M., et al., *Comparison of Proteomic Assessment Methods in Multiple Cohort Studies*. Proteomics, 2020. **20**(12): p. e1900278.
- 848 62. Anderson, J.R., et al., *Optimization of Synovial Fluid Collection and Processing for NMR Metabolomics and LC-MS/MS Proteomics*. J Proteome Res, 2020. **19**(7): p. 2585-2597.
- 849 63. Ali, N., et al., *Proteomics Profiling of Human Synovial Fluid Suggests Increased Protein Interplay in Early-Osteoarthritis (OA) That Is Lost in Late-Stage OA*. Mol Cell Proteomics, 2022. **21**(3): p. 100200.
- 850 64. Struglics, A., et al., *Technical performance of a proximity extension assay inflammation biomarker panel with synovial fluid*. Osteoarthr Cartil Open, 2022. **4**(3): p. 100293.
- 851
- 852
- 853
- 854
- 855
- 856
- 857
- 858
- 859
- 860
- 861
- 862
- 863
- 864
- 865
- 866
- 867
- 868
- 869
- 870
- 871
- 872
- 873
- 874
- 875
- 876
- 877

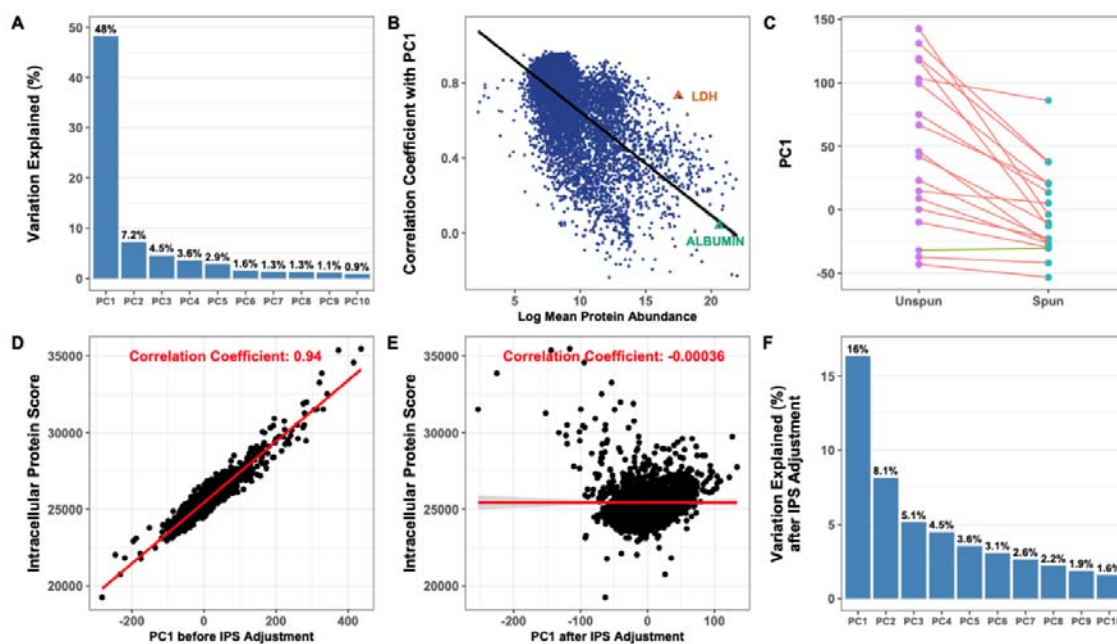
878 65. Struglics, A., et al., *Changes in Cytokines and Aggrecan ARGS Neopeptide in Synovial Fluid*  
879 *and Serum and in C-Terminal Crosslinking Telopeptide of Type II Collagen and N-Terminal*  
880 *Crosslinking Telopeptide of Type I Collagen in Urine Over Five Years After Anterior Cruciate*  
881 *Ligament Rupture: An Exploratory Analysis in the Knee Anterior Cruciate Ligament,*  
882 *Nonsurgical Versus Surgical Treatment Trial.* *Arthritis & Rheumatology*, 2015. **67**(7): p. 1816-  
883 1825.

884

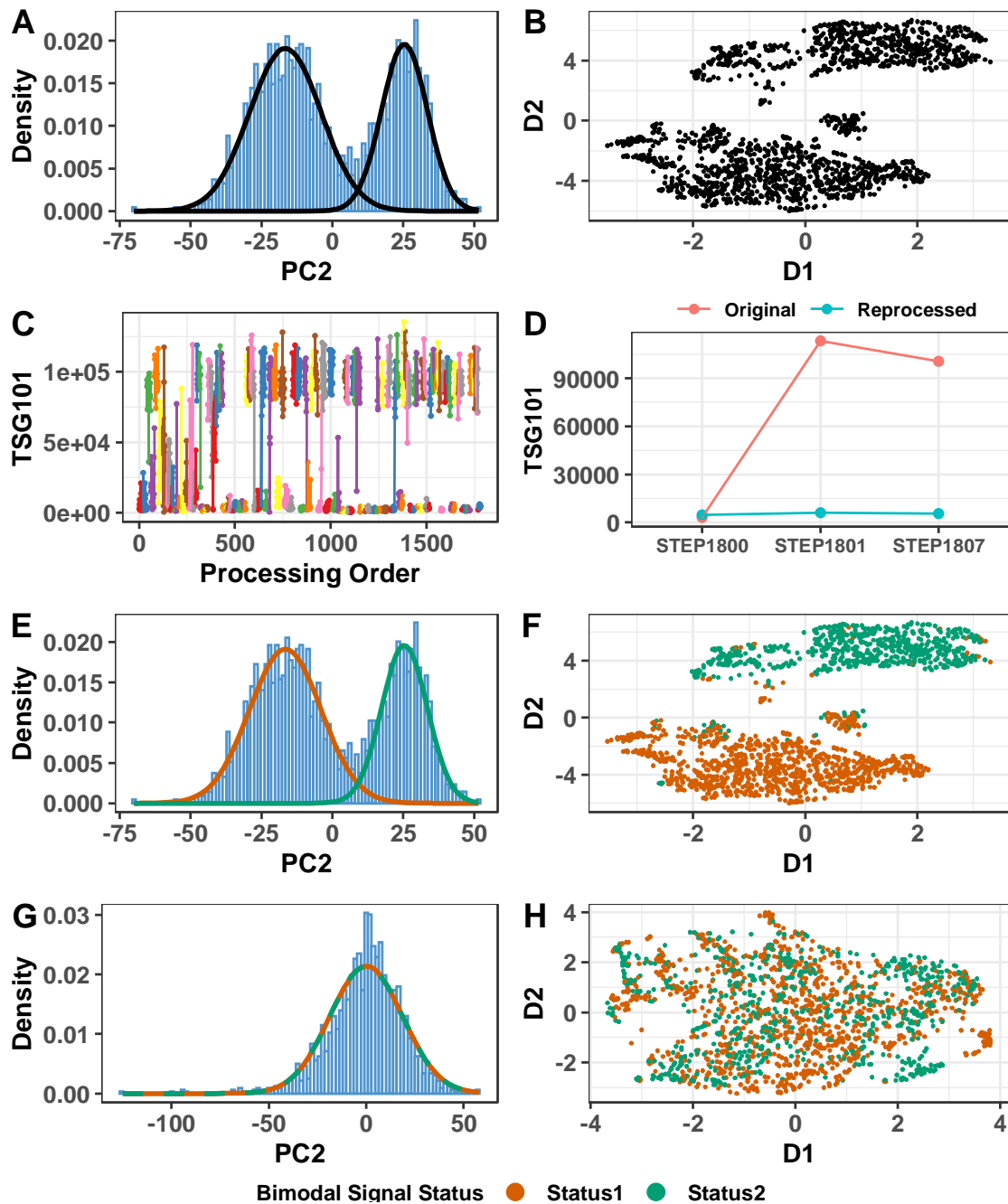


**Figure 1.** Assessment of the effects of each standardisation step on (A) mean %CV and (B) mean  $R^2$  across all proteins for pooled sample replicates, stratified by OA and acute knee injury respectively. Assessment of Pearson correlation coefficients between protein expression in samples measured by the SOMAscan platform and by prior immunoassay for nine select proteins across normalisation steps for (C) OA and (D) acute knee injury. The normalisation steps included hybridisation normalisation (HN), plate scaling using plasma calibrators (PS), median signal normalisation (MN) and calibration using plasma calibrators (PC). Correlation between the RFUs (SOMAscan) and absolute concentrations for the nine proteins across the two methods are shown. RFUs, relative fluorescence units; %CV, % coefficient of variation; osteoarthritis, OA; Activin A, Inhibin beta A chain; FGF2, Fibroblast growth factor 2; IL6, Interleukin-6; IL8, Interleukin-8; MCP1, C-C motif chemokine 2; MMP3, Stromelysin-1; TGFβ1, Transforming growth factor beta-1; TIMP1, Tissue inhibitor of metalloproteinase 1; TSG6, Tumor necrosis factor-inducible gene 6.



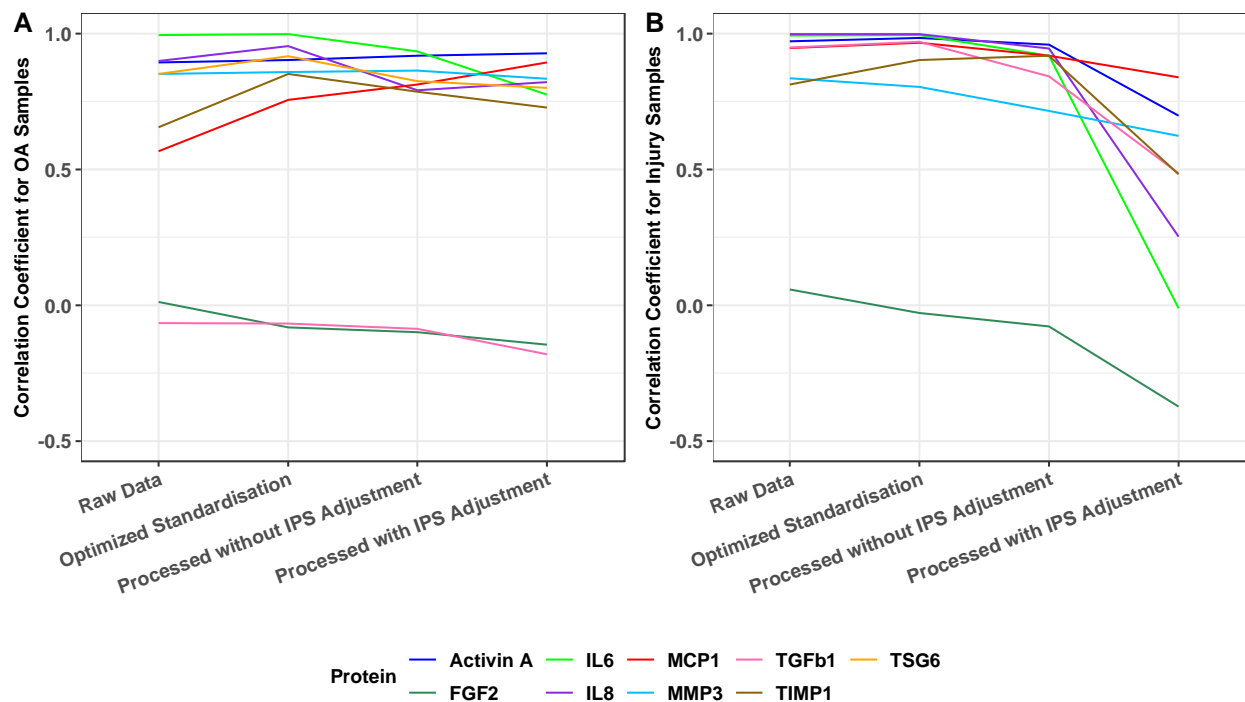


**Figure 2.** (A) Variation explained (%) by the top 10 PCs derived from the standardised log abundance proteomic data. (B) Correlation between PC1 and protein abundance, with two high-abundance proteins (albumin, a soluble serum protein, and LDH, an intracellular protein) marked. Protein abundance is calculated as the standardized RFU for each protein adjusted by the protein's dilution factor used in the SomaScan assay (the "dilution bin"). (C) Comparison of variation explained (%) by PC1 between 18 pairs of SF samples that were centrifuged (spun) or not (unspun) after aspiration and prior to freezing, with paired samples from the same participant joined by separate lines. Red lines show samples that had an increased PC1 prior to spinning, and the green line where it was decreased. Correlation between PC1 and intracellular protein score (D) before and (E) after IPS adjustment. (F) Variation explained by the top 10 PCs derived from the batch corrected and IPS adjusted log abundance proteomic data. In all cases, correlation is measured using the Pearson correlation coefficient. IPS, Intracellular Protein Score; PC, principal component; LDH, Lactate dehydrogenase.

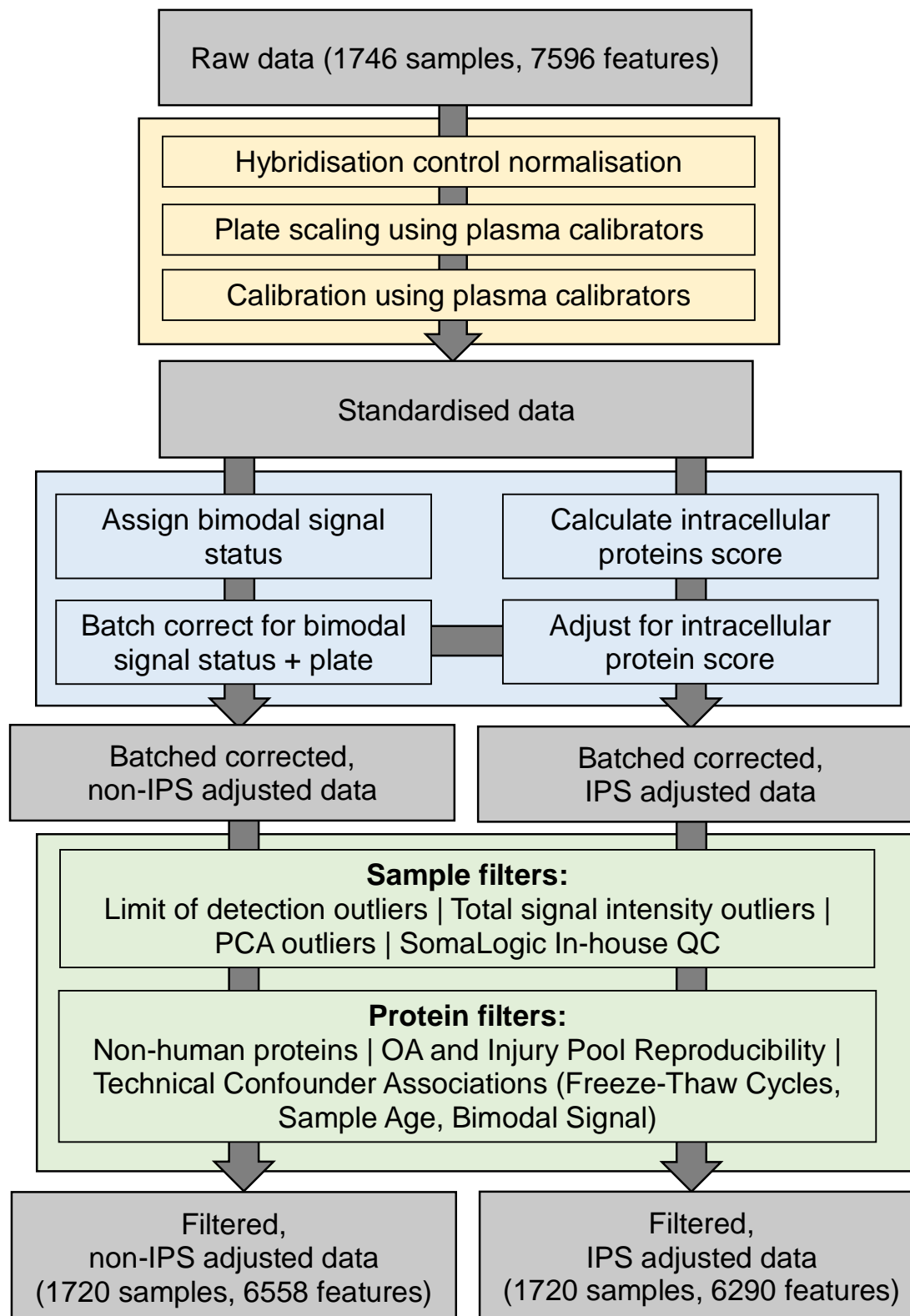


**Figure 3.** (A) Distribution of the second principal component (PC2) derived from the standardised log abundance data, showing a bimodal distribution. (B) UMAP visualisation of two reduced dimensions (D1 and D2) of the top PCs of the standardised log abundance data. (C) Example of a strongly bimodal protein measurement, TSG101, RFU (y-axis) against Oxford laboratory processing order (x-axis) and coloured by laboratory processing batch (with only points within the same processing batch

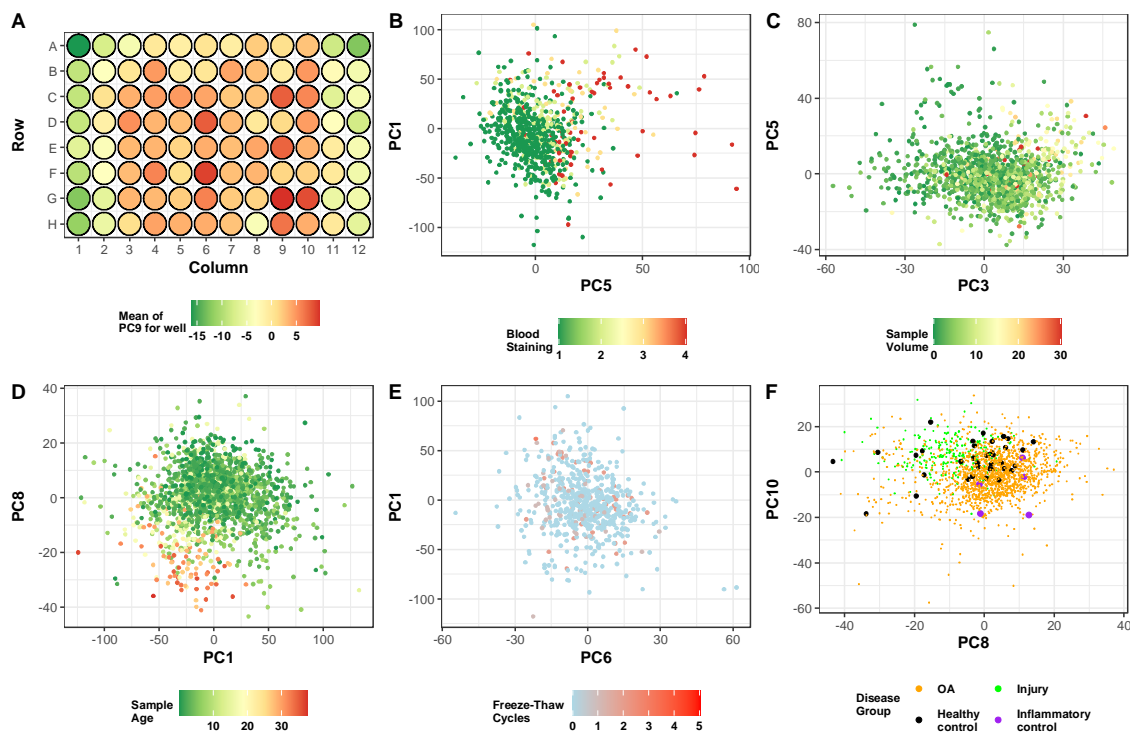
connected by lines). Note that the ‘flipping’ between high and low signal status occurred primarily when processing batch changed, and only rarely within processing batch. This effect was particularly strong among sample batches that were processed later in processing order. (D) The same example protein measurement for three independent SF samples before (original) and after they were re-processed and re-assayed, showing that bimodal status changed after laboratory re-processing. (E) Distribution of PC2 derived from standardised log abundance data, showing the two probability density functions of the Gaussian Mixture Model used to classify samples into the two bimodal signal status groups. (F) UMAP visualisation of two reduced dimensions (D1 and D2) of the top PCs of the standardised log abundance data, colored by the inferred bimodal signal status. (G) Histogram of PC2 of the batch corrected log abundance data, with the now near-identical distributions of the two bimodal signal status groups shown as colored lines, (H) UMAP visualisation on two reduced dimensions (D1 and D2) of the top PCs of the batch corrected log abundance data, colored by the inferred bimodal signal status. RFUs, relative fluorescence units; PC, Principal Component; TSG101, Tumor susceptibility gene 101 protein; UMAP, Uniform Manifold Approximation and Projection.



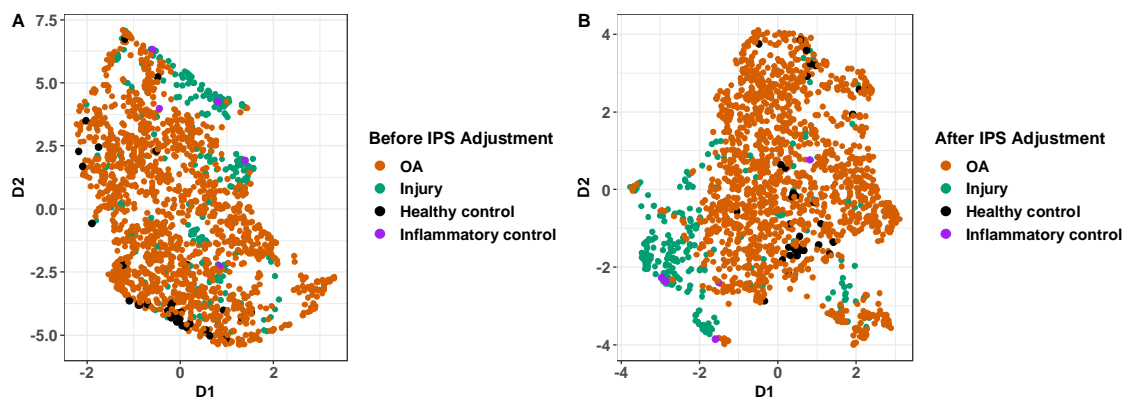
**Figure 4.** Correlation between SOMAscan relative frequency abundance (RFU) and abundance measured using orthogonal immunoassays for 9 selected proteins at different stages of SOMAscan data processing, for (A) osteoarthritis and (B) acute knee injury samples. Correlation was measured using the Pearson correlation coefficient. Raw data refers to the raw RFUs without any processing, optimised standardisation was the data standardised using our selected optimal normalization steps (Figure 1), processed without IPS adjustment refers to data that has been batch corrected for bimodal signal status and plate but not IPS adjusted, and processed with IPS adjustment refers to samples that have undergone both batch correction and IPS adjustment. IPS, Intracellular Protein Score; Protein name abbreviations as in Figure 1.



**Figure 5.** Overview of the final data processing and quality control pipeline for synovial fluid SOMAscan data used by the STEpUP OA consortium, broken down into three stages: standardisation (yellow box), technical confounder correction (blue box) and filtering (green box). More details on filtering thresholds, and the number removed by each filter, can be found in Supplementary Table S4.



**Figure 6.** Visualisation of selected predefined confounders against select principal components of the batch corrected, filtered, IPS adjusted data. (A) The average value of PC9 (most strongly associated with plate position) by sample well position, (B-F) visualisation of the two PCs most strongly associated with each confounder, coloured by confounder value. Pre-defined confounders shown are (B) blood staining grade of sample after aspiration assessed by visual inspection, (C) volume of sample taken during aspiration, (D) age of the sample in years, measured from aspiration to sample processing at Oxford, (E) the number of times the sample was thawed and re-frozen before sample processing at Oxford, (F) the disease group of the sample (osteoarthritis [OA], acute knee injury [Injury], healthy control, inflammatory arthritis control).



**Figure 7.** UMAP visualisation of two reduced dimensions (D1 and D2) of the top PCs of the log abundance data with (A) and without (B) IPS adjustment followed by filtering, coloured by disease group. These groups were osteoarthritis (OA, acute knee injury (injury), healthy controls, inflammatory arthritis controls.

UMAP, Uniform Manifold Approximation and Projection.