EXTENDED REPORT

Establishment of reference intervals for osteoarthritis-related soluble biomarkers: the FNIH/OARSI OA Biomarkers Consortium

Virginia B Kraus,[1] David E Hargrove,[2] David J Hunter,[3] Jordan B Renner,[4] Joanne M Jordan[5]

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-209253).

For numbered affiliations see end of article.

Correspondence to
Dr Virginia Byers Kraus, Duke Molecular Physiology Institute, Box 104775, 300 N Duke St, Durham, NC 27701, USA; vbk@duke.edu

Received 23 January 2016
Revised 4 May 2016
Accepted 17 May 2016
Published Online First 24 June 2016

▶ http://dx.doi.org/10.1136/annrheumdis-2016-209252

To cite: Kraus VB, Hargrove DE, Hunter DJ, *et al. Ann Rheum Dis* 2017;**76**:179–185.

ABSTRACT

Objective To establish reference intervals for osteoarthritis (OA)-related biomarkers used in the Foundation for the National Institutes of Health (FNIH) OA Biomarkers Consortium Project.

Methods A total of 129 'multijoint controls' were selected from 2722 African-American and Caucasian men and women in the Johnston County Osteoarthritis Project. The majority (79%) of those eligible (with biospecimens and baseline data) also had one or more follow-up evaluations 5–15 years later. Multijoint controls were selected to be free of radiographic hand, hip, knee and lumbar spine osteoarthritis (OA), to have no knee or hip symptoms, and minimal hand and spine symptoms at all available time points. Eighteen biomarkers were evaluated in serum (s) and/or urine (u) by ELISA. Reference intervals and partitioning by gender and race were performed with EP Evaluator software.

Results Controls were 64% women, 33% African-Americans, mean age 59 years and mean body mass index 29 kg/m^2. Three biomarkers were associated with age: sHyaluronan (positively), sN-terminal propeptide of collagen IIA (positively) and sCol2-3/4 C-terminal cleavage product of types I and II collagen (negatively). Exploratory analyses suggested that separate reference intervals may be warranted on the basis of gender for uC-terminal cross-linked telopeptide of type II collagen (uCTXII), sMatrix metalloproteinase-3, uNitrated type II collagen degradation fragment (uCol2-1 NO2) and sHyaluronan, and on the basis of race for uCTXII, sCartilage oligomeric matrix protein, sC-terminal cross-linked telopeptide of type I collagen and uCol2-1 NO2.

Conclusions To our knowledge, this represents the best and most stringent control group ever assayed for OA-related biomarkers. These well-phenotyped controls, representing a similar age demographic to that of the OA Initiative-FNIH main study sample, provide a context for interpretation of OA subject biomarker data. The freely available data set also provides a reference for future human studies.

INTRODUCTION

A reference interval (RI) is the central 95% range—or normal range—for endogenous analytes of a healthy person. Traditionally, reference ranges for biomarkers are established using commercially purchased 'normal' samples from 'normal' blood donors. Such 'normals' are rarely adequate controls for osteoarthritis (OA) studies as they are rarely if ever ascertained for OA status or from an appropriate age group. Given the very high prevalence of OA with ageing,[1 2] lack of OA phenotyping could result in a reference 'normal' population with as much as 18% prevalence of OA (for individuals over 60 years of age)[3] and up to 75% prevalence (based on the prevalence of hand OA in older age groups).[2] The objective of this ancillary study was to establish RIs for OA-related biomarkers using biospecimens from stringently phenotyped age-appropriate controls and to aid in the interpretation of results from the Foundation for the National Institutes of Health/Osteoarthritis Research Society International (FNIH/OARSI) OA Biomarkers Consortium Project.[4] We hypothesised that the RI ranges for these 'multijoint controls' will facilitate future formal Food and Drug Administration (FDA) qualification of an FNIH/OARSI OA-related biomarker panel for OA prognosis and for monitoring the efficacy of interventions aimed at slowing or stopping OA progression. Our long-term objective is to pursue US FDA qualification of OA-related biomarkers for their use as drug development tools.

METHODS

Selection of 'multijoint controls'

The total sample (N=2722) from which multijoint controls were selected was from the Johnston County OA Project (JoCoOA), a population-based study of OA in a single county in North Carolina, USA; details of the sampling and recruitment strategy have been previously reported.[5] Eligible participants (N=1930) for this substudy had baseline biospecimens and radiographic and symptom data: 1518 (79%) of these eligible participants also had data from one or more follow-up time points up to 5–15 years after baseline; 412 participants had baseline data only. To identify multijoint controls, participants were selected who had no radiographic evidence of OA in either knee (OA defined as Kellgren-Lawrence (KL) grade ≥1) or either hip (OA defined as KL grade ≥2) at any available time point, and no knee or hip symptoms (pain, aching or stiffness on most days) at any available time point (figure 1). Further, based on definitions used in the Genetics of Generalized Osteoarthritis (GOGO) study,[6] multijoint controls had no radiographic hand OA (OA defined as involvement of at least three joints with radiographic KL2 OA distributed bilaterally and involving at least one distal interphalangeal joint), or radiographic lumbar spine

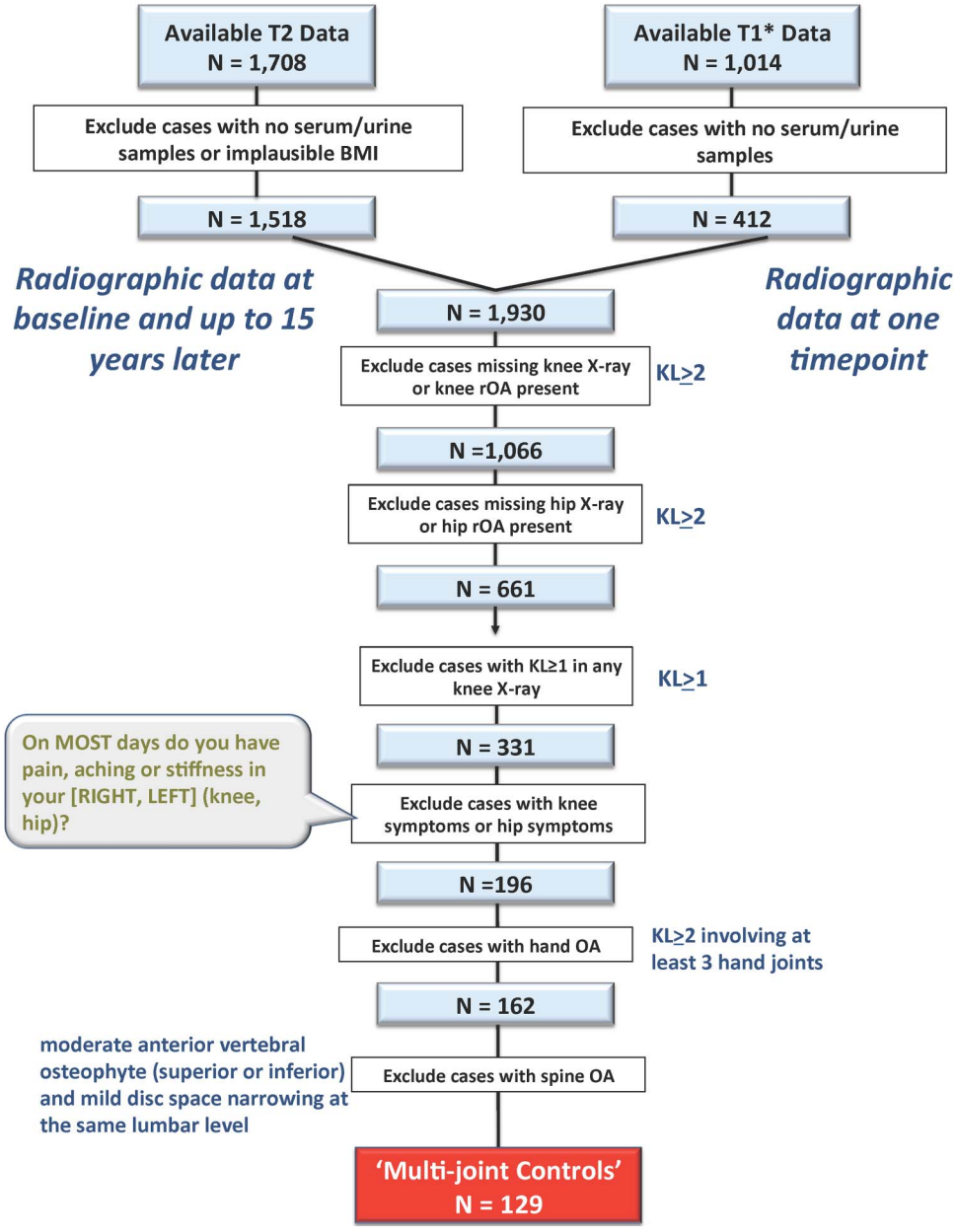

Figure 1 Algorithm for selecting the non-OA 'multijoint control' subjects. From a total sample of 2722 individuals, a total of 129 controls were selected. The basis for selection included availability of biospecimens, radiographic and clinical data, preferably with at least one 5-year follow-up examination (available for 79% of eligible subjects), no radiographic evidence of knee OA (knees KL0 bilaterally) or hip OA (hips KL0–1 bilaterally) at any available time point, and no knee or hip symptoms (no pain, aching or stiffness on most days) at any available time point. Multijoint controls also had minimal radiographic hand OA or lumbar spine radiographic OA at any available time point. BMI, body mass index; KL, Kellgren-Lawrence; OA, osteoarthritis. T1* is Time 1 (baseline) for the second wave of recruited subjects to JoCo, T2 is Time 2 representing the first follow up (1999–2004).

OA (OA defined as disc narrowing grade ≥1 and anterior osteophyte >1 at same level) at any available time point. This selection process yielded a total of n=129 multijoint controls participants with minimal or no radiographic burden of disease, two-thirds of whom (n=83) had available data to confirm lack of OA development over the subsequent 5–15 years of observation. All participants provided data on pain of the lower extremities (based on the Western Ontario and McMaster Universities Arthritis Index (WOMAC)) and depressive symptoms (using the Center for Epidemiological Studies-Depression 20 question standardised depressive symptoms assessment scale).[7 8]

Radiographic imaging

Radiographs were obtained of the knees, hips, hands and spine as described previously.[9] All radiographs were read by a single experienced musculoskeletal radiologist (JBR), previously shown to have high intrarater and inter-rater reliability (kappa 0.89 and 0.86, respectively).[10]

Biospecimen collection

The majority of sera were obtained fasting (ie, no food for at least the prior 8 hours, 11.6%) or 2 hours postprandially (52.7%). All urine (cell-free supernatants) were collected as second morning void specimens. Specimens were collected at

the time of the research clinic visit, stored on ice and frozen within 8 hours of collection at −20°C, then transferred within a few weeks to long-term storage at −86°C. The biospecimens used for this research had never been thawed prior to these analyses.

Biomarker assays

The biomarkers selected for this study were the ones chosen for the main FNIH OA Biomarkers Consortium Project.[4] These biomarkers were agreed on through consensus of the OARSI/FDA Biomarkers Working Group.[11] Their selection was based on availability of commercially available kits and evidence that each biomarker met criteria for one or more of the burden of disease, investigation, prognostic, efficacy of intervention, diagnostic and safety biomarker categories.[11] The chosen biomarkers included analytes reflecting catabolic and anabolic processes of cartilage and bone.

Several biomarkers were quantified in both serum (s) and urine (u), resulting in a total of 18 separate biomarkers assayed (note, the same analyte in different body fluids is referred to as a separate biomarker for clarity). Biomarkers analysed in serum alone were cartilage oligomeric matrix protein (COMP), hyaluronan (HA), C-propeptide of type II collagen (CPII), N-terminal propeptide of collagen IIA (PIIANP), chondroitin sulfate 846 epitope (CS846), the C-terminal cross-linked telopeptide of type I collagen (CTXI) and matrix metalloproteinase-3 (MMP-3). Biomarkers analysed in serum and urine were Col2-3/4 C-terminal cleavage product of types I and II collagen (C1,2C), Col2-3/4 C-terminal cleavage product of type II collagen (C2C in serum, C2C HUSA in urine), nitrated epitope of the α-helical region of type II collagen (Coll2-1 NO2) and the cross-linked N-telopeptide of type I collagen (NTXI). Biomarkers analysed in urine alone were the C-terminal cross-linked telopeptide of type II collagen (CTXII), and α and β isomerised versions of the C-terminal cross-linked telopeptide of type I collagen (CTXIα and CTXIβ). All urine biomarkers were normalised to creatinine. The kit manufacturers, catalogue numbers and reported lower limit of quantification (LLOQ) for each biomarker are listed in table 1. Kits with the same lot number were used for all sample analyses of each particular biomarker; the lots used for this research, however, were different from those used in the main FNIH OA biomarker study.

All biomarker assays were performed by LabCorp Clinical Trials, a Clinical Laboratory Improvement Amendments (CLIA) and College of American Pathologists (CAP) certified division within LabCorp by personnel blinded to the clinical data. All samples were analysed in duplicate. Interassay coefficients of variation (CVs, provided in table 1) are based on the average CVs of high and low concentration control samples provided with each kit and run on each plate; when available (as for serum MMP-3 and HA), a third medium concentration standard was also run on each plate and combined with the high and low concentrations for an overall CV. The initial dilutions for sample measurements (listed in online supplementary table S1) were agreed upon by the kit manufacturers. Samples with concentrations above the highest standard were repeated at a higher dilution until results were within the linear range of the assay; thus all high values were quantifiable. When biomarker results were below the lowest standard, the kit manufacturer was consulted to determine lower dilutions that could be tested without the likelihood of incurring problems with assay inhibition; lower (1:2) dilutions were deemed feasible for three markers including sMMP-3 (6 samples), uCTXIα (43 samples) and uCTXIβ (13 samples). For many of the biomarkers, there were still some samples with values below the LLOQ; these concentrations were imputed the same way as for the main FNIH OA Biomarker Project, that is, by interpolation from the standard curve extended from the lowest standard to zero. These biomarker data are freely available on the OA Initiative (OAI) website (https://oai.epi-ucsf.org/datarelease/FNIH.asp).

Table 1 Subject characteristics

Clinical parameter	Category	Value
Age in years, mean (SD)		59.4 (8.2)
Age groups	50 to <55	33 (26%)
	55 to <65	78 (61%)
	65 to <75	10 (8%)
	75+	8 (6%)
BMI in kg/m^2, mean (SD)		29.1 (5.9)
BMI groups, n (%)	≤30	85 (66%)
	>30	44 (34%)
Caucasian, n (%)		85 (66%)
Women, n (%)		82 (64%)
CES-D, mean (SD)		5.7 (7.2)*
WOMAC pain, mean (SD)	0–50	1.0 (2.5)
WOMAC pain groups, n (%)	=0	98 (76%)
	≥1	31 (24%)
Symptoms, n (%)	Knee	0 (0%)
	Hip	0 (0%)
	Hand	24 (19%)
	Foot	22 (18%)
	Low back	33 (26%)

*CES-D <16 means no clinically significant depression.
BMI, body mass index; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

Statistical analysis and determination of RIs

The RIs were determined with the RIs establishment module of EP Evaluator (Data Innovations) using the transformed parametric method with the exception of uC1,2C that used the non-parametric method because transformation did not improve the Gaussian distribution for this biomarker. The transformed parametric method first attempts to change the scale of the data so it has a Gaussian distribution. The software computes the central 95% RI (lower and upper limits) and the 90% CIs around these limits. The software also provides the confidence ratio (CR) defined as the ratio of the average CI width to the RI width. The primary determinant of the CR is sample size with smaller (improved) CR as sample size increases. A CR value of ≤0.10 is desirable; values of over 0.30 are flagged in the report. The EP Evaluator partitioning test was used to determine if separate RIs might be justified for population subclasses based on gender or ethnicity (tested separately). When the SD ratio, calculated as the larger SD divided by smaller SD of the two subgroups (for instance male vs female and African-American vs Caucasian) is >1.5, separate RIs may be warranted. Partitioning with EP Evaluator based on gender and race was exploratory as the recommended group size of 60 was not always available for each subclass of subjects. Spearman correlations were used to assess for a significant associations ($p<0.05$) between the biomarker concentrations and age and body mass index (BMI). The biomarker concentrations of multijoint controls and the OA subjects of the FNIH main study were compared in two ways. First, we tabulated the proportion of subjects in the main FNIH study

with biomarker concentrations greater than or equal to the median of the multijoint controls for catabolic and inflammatory biomarkers, and less than or equal to the median of the multijoint controls for anabolic biomarkers. Second, we compared the median biomarker concentrations of the multijoint controls with the OA cases (n=194 with joint space loss progression and pain worsening over 48 months) and OA comparators (n=406 who did not meet the OA case definition, ie, did not have the combination joint space loss progression and pain worsening over 48 months, instead had pain worsening only (n=103), joint space loss progression only (n=103) or neither (n=200) over 48 months) in the FNIH main study. The non-parametric rank-sum test was used to evaluate differences in biomarker concentrations between the 46 multijoint controls without follow-up and the 83 multijoint controls with follow-up time-point evaluations.

RESULTS

These multijoint controls were selected from the JoCo population on the basis of the criteria outlined in figure 1 including the lack of radiographic OA of the knee, hip, hands and spine as well as the lack of knee and hip symptoms. A minority of this cohort (<26%) had any symptoms in the hands, feet or lumbar spine (table 1). The one sense in which these controls could be considered at risk for OA was their age, since by design, they were also selected to match the age demographic (≥45 years old) typical for individuals with OA and the main FNIH OA biomarker study. The sample of 129 multijoint controls consisted of 64% women, 33% African-Americans, mean age 59 years and mean BMI 29 kg/m^2 (table 1). The majority of subjects (76%) had WOMAC scores of 0. In addition, this cohort had very low clinical depressive symptoms scores (mean score 5.7) indicating no evidence of clinically significant depression based on an established threshold for depression for scores ≥16 (http://cesd-r.com/cesdr/).

Three biomarkers correlated significantly with age: sHA (positively), sPIIANP (positively) and sC1,2C (negatively). No biomarkers were correlated with BMI (table 2). The central 95% CIs for the 18 OA-related biomarkers are provided in table 2. The CR for each of these biomarkers was below the generally acceptable threshold of 0.3.

A comparison of the median biomarker concentrations of the multijoint controls with the FNIH OA biomarker main study results[4] showed that for six catabolic or proinflammatory biomarkers (sCTXI, sCol2-1 NO2, sMMP-3, sHA, uC2C and uCTXII), more than half of the total subjects in the main FNIH OA Biomarkers Consortium study (all knee OA subjects) had concentrations greater than or equal to the median for the multijoint controls (see online supplementary table S1); for all three anabolic biomarkers (sCPII, sPIIANP and sCS846), more than half of the total subjects in the main FNIH study had concentrations less than or equal to the median for the multijoint controls. We also compared median values of the multijoint controls with the median values in the FNIH OA progressor cases and OA comparators. The median biomarker concentrations for both OA cases and OA comparators from the main FNIH study were above the medians for multijoint controls for the catabolic and proinflammatory biomarkers, sCTXI, sCol2-1 NO2, sHA, sMMP-3, uC2C and uCTXII; the median biomarker concentrations for OA cases only were above the medians for multijoint controls for uNTXI and uCTXIß. The median biomarker concentrations for both OA cases and OA comparators from the main FNIH study were below the median for multijoint controls for the anabolic biomarkers, sCS846, sCPII and sPIIANP.

Results of partitioning with EP Evaluator (table 3) suggested that separate RIs may be warranted on the basis of gender for uCTXII, sMMP-3, uCol2-1 NO2 and sHA (listed from most to least gender variation based on SD ratio). Compared with men, women had higher mean concentrations of uCTXII and sCol2-1 NO2 but lower mean concentrations of sHA and sMMP3. Separate RIs may also be warranted on the basis of race for uCTXII, sCOMP, sCTXI and uCol2-1 NO2 (listed from most to least race variation based on SD ratio). Compared with Caucasians, African-Americans had higher mean concentrations of uCTXII, sCOMP and sCTXI but lower mean concentrations of uCol2-1 NO2.

The mean concentration of several biomarkers differed significantly between the subgroup of multijoint controls without and with follow-up time points: sC2C, sCol2-1 NO2, sHA, sPIIANP, uCol2-1 NO2, uC1,2C and uCTXIα (see online supplementary table S2).

CONCLUSIONS

We refer to this cohort as multijoint controls because they had little or no radiographic OA and little or no joint pain for the major joint groups throughout the body. To our knowledge, this represents the best and most stringent control group ever assayed for OA-related biomarkers. These well-phenotyped controls represent a similar age demographic to that of the OAI-FNIH main study sample so they should provide an optimal reference control for the main study. These results and freely available data set also provide a reference for future human OA studies.

Based on the comparison of median biomarker concentrations of non-OA multijoint controls and the knee OA main FNIH study subjects, there was considerable overlap of concentration values in the two groups. Nevertheless, several biomarkers emerged as showing a different distribution of concentrations, with catabolic and proinflammatory biomarkers in FNIH OA subjects above the multijoint control median, and anabolic biomarkers in FNIH OA subjects below the multijoint control median.

Results of this study suggest age associations for three biomarkers (sHA, sPIIANP and uC1,2C). Our results are concordant with previous reports of positive associations with age for sHA[12] and sPIIANP.[13] Based on a SD ratio >1.5, separate RIs may be warranted on the basis of gender for four biomarkers (sHA, sMMP3, uCol2-1 NO2 and uCTXII) and on the basis of race for four biomarkers (sCTXI, sCOMP, uCol2-1 NO2 and uCTXII). The greater variance in women underlies the suggestion for separate RIs for sHA and uCol2-1 NO2, while the greater variance in men underlies the suggestion for separate RIs for sMMP3 and uCTXII. In the case of race, the greater variance in African–Americans in uCTXII, sCOMP and sCTXI underlies the suggestion for separate RIs of these biomarkers. The influences responsible for these differences in means and variances are generally not known although presumably, genetics and hormonal influences are likely playing a role.

The number of females and Caucasians in this sample exceeded the Clinical and Laboratory Standards Institute (CLSI)[14] recommended 60 reference subjects for initial determination of whether a single combined RI is suitable or separate RIs are justified. The CLSI guideline further recommended collecting an additional 60 samples from each subclass for confirmation of separate RIs if the initial statistical analyses indicated significant subclass differences. Thus, additional independent multijoint control biomarker analyses would be a welcome addition to the

Table 2 RI estimations for whole cohort (N=129)

Biomarker	Manufacturer	Mean (SD) Median	Range	Central 95% RI Lower limit (90% CI)	Central 95% RI Upper limit (90% CI)	Confidence ratio	Significant Spearman correlations Age/BMI
Serum C1,2C (ng/mL)	IBEX	407.13 (107.43) 400.00	180 to 720	228 (211 to 245)	646 (609 to 684)	0.13	p>0.05
Serum C2C (ng/mL)	IBEX	278.23 (48.98) 271.0	186 to 433	196 (188 to 205)	383 (367 to 400)	0.13	p>0.05
Serum Col2-1† NO2 (nM)	Artialis	9.01 (4.42) 8.21	2.24 to 27.98	3.24 (2.89 to 3.64)	20.20 (17.99 to 22.68)	0.16	p>0.05
Serum CPII (ng/mL)	IBEX	1450.90 (401.08) 1379.0	649 to 3074	830 (777 to 887)	2362 (2210 to 2523)	0.14	p>0.05
Serum CS846 (ng/mL)	IBEX	95.2 (42.0) 86.0	26 to 129	39 (35 to 43)	197 (177 to 218)	0.15	p>0.05
Serum CTXI† (ng/mL)	IDS	0.31 (0.22) 0.24	0.06 to 1.61	0.079 (0.068 to 0.091)	0.808 (0.697 to 0.936)	0.18	p>0.05
Serum COMP (ng/mL)	Biovendor	1016.12 (650.19) 898.0	403 to 5797	404 (364 to 448)	2070 (1867 to 2296)	0.15	p>0.05
Serum HA† (ng/mL)	Corgenix	36.1 (29.4) 29.0	6 to 239	8 (7 to 9)	104 (89 to 123)	0.19	ρ 0.19 (p=0.03) with age
Serum MMP-3† (ng/mL)	Invitrogen	16.97 (16.06) 13.21	3.22 to 143.64	3.72 (3.17 to 4.37)	47.89 (40.74 to 56.30)	0.19	p>0.05
Serum NTXI (nM BCE)	Ostermark/ Alere	16.02 (5.06) 15.0	7 to 40	9 (8 to 9)	27 (25 to 29)	0.14	p>0.05
Serum PIIANP (ng/mL)	EMD Millipore	2879.90 (705.58) 2718.0	1068 to 4695	1690 (1577 to 1810)	4431 (4195 to 4676)	0.13	ρ 0.27 (p=0.002) with age
Urine Col2-1 N02 (nmol/mmol Cr)	Artialis	0.0274 (0.0234) 0.0235	0 to 0.1388	0 (0 to 0.0004)	0.086 (0.071 to 0.091)	0.15	p>0.05
Urine C1,2C* (ng/mmol Cr)	IBEX	0.011 (0.010) 0.010	0 to 0.05	0.0 (0 to 0)	0.04 (0.03 to 0.05)	0.27	ρ −0.20 (p=0.02) with age
Urine C2C† (HUSA) (ng/mmol Cr)	IBEX	100.2 (67.8) 90.0	0 to 427	6 (2 to 10)	268 (239 to 297)	0.13	p>0.05
Urinary CTXII† (ng/mmol Cr)	IDS	264.28 (520.80) 175.00	49.3 to 5795.53	50.46 (42.76 to 59.56)	691.34 (585.78 to 815.92)	0.19	p>0.05
Urine NTXI‡ (nmol BCE/mmol Cr)	Ostermark/ Alere	36.93 (21.37) 32.0	10 to 148	11 (10 to 13)	90 (79 to 103)	0.17	p>0.05
Urine (α) CTXI (μg/mmol Cr)	IDS	0.57 (0.55) 0.42	0 to 4.18	0.04 (0.03 to 0.06)	1.81 (1.56 to 2.10)	0.16	p>0.05
Urine (β)‡ CTXI (μg/mmol Cr)	IDS	2.51 (1.88) 1.99	0.21 to 11.08	0.46 (0.38 to 0.55)	8.20 (6.83 to 9.85)	0.21	p>0.05

Significant Spearman Correlations were defined as p>0.05.
All urinary biomarkers are normalised to mmol/L urinary creatinine (Quidel).
Confidence ratio is the average CI width to the RI width; a value ≤0.10 is desirable, values <0.30 are considered acceptable.
*Transformed parametric method to compute RI with the exception of uC1,2C that used the non-parametric method.
†Median concentrations of cases and controls in the main FNIH study were above the medians of multijoint controls for: sCTXI, sHA, sMMP-3, uC2C, uCTXII.
‡Median concentrations of cases in the main FNIH study were above the medians of multijoint controls for: uNTXI and uCTXIß.
BCE, bone collagen equivalents; BMI, body mass index; C1,2C, C-terminal neoepitope of 3/4 piece generated by cleavage of types I and II collagen by collagenases; C2C, C-terminal neoepitope of 3/4 piece generated by cleavage of type II collagen by collagenases; Col2-1 NO2, nitrated type II collagen degradation fragment; COMP, cartilage oligomeric matrix protein; CPII, type II collagen carboxy propeptide cleaved following release of newly synthesised procollagen into matrix; Cr, creatinine; CS846, chondroitin sulfate epitope of aggrecan; CTXI, C-terminal telopeptide of type I collagen (also called CrossLaps), α CTXI contains the newly synthesised α form of aspartic acid, β CTXI contains the age-related β form of aspartic acid; CTXII, C-terminal telopeptide of type II collagen (also called CartiLaps); FNIH, Foundation for the National Institutes of Health; HA, hyaluronan; MMP-3, matrix metalloproteinase-3 also called stromelysin-1; NTXI, N-terminal telopeptide of type I collagen; PIIANP, N-propeptide encoded by exon 2 of type II collagen; RI, reference interval.

field although, as demonstrated here, require a large initial cohort to achieve the requisite sample size for confirmation.

A great strength of the study was the long period of follow-up for the majority of the multijoint controls, verifying their lack of OA development well beyond the time of blood and urine sampling. Interestingly, the mean concentrations of a few of the biomarkers, sCol2-1 NO2, sPIIANP, uC1,2C and uCTXIα, varied in a manner consistent with the possibility of the presence of early but radiographically negative OA in some participants in the subgroup without follow-up time points. This suggests that these biomarkers would be worth investigating further as potential biomarkers for predicting incident radiographic OA. These data also suggest that the differences in concentrations between OA subjects and OA-free controls for these biomarkers could be underestimated in these analyses.

In conclusion, the ability to diagnose and prognose OA with biomarkers is dependent on a clear understanding of normal RIs. For a highly prevalent and heterogeneous disorder such as

Table 3 Partitioning of biomarker concentrations by gender and race

Biomarker	Gender	Mean	SD	Central 95% RI		SD ratio	Race	Mean	SD	Central 95% RI		SD ratio
				Lower limit	Upper limit					Lower limit	Upper limit	
Serum C1,2C (ng/mL)	All	407.1	107.4	228	646	1.3	All	407.1	107.4	228	646	1.0
	Female	417.4	97.2	253	629		AA	430.5	115.4	241	683	
	Male	389.1	122.3	194	663		Cauc	395.1	101.7	223	624	
Serum C2C (ng/mL)	All	278.2	49.0	196	398	1.0	All	278.2	49.0	196	383	1.0
	Female	284.5	49.0	201	390		AA	284.1	48.5	203	387	
	Male	267.3	47.6	189	367		Cauc	275.2	49.2	193	381	
Serum Col2-1 NO2 (nM)	All	9.01	4.42	3.24	20.20	1.5	All	9.01	4.42	3.24	20.20	1.0
	Female	9.87	4.81	3.77	21.19		AA	9.66	4.33	3.65	21.09	
	Male	7.51	3.26	2.74	16.96		Cauc	8.68	4.46	3.07	19.63	
Serum CPII (ng/mL)	All	1450.9	401.1	830	2362	1.1	All	1450.9	401.1	830	2362	1.1
	Female	1510.0	376.5	918	2344		AA	1564.3	371.2	981	2371	
	Male	1347.7	425.4	725	2298		Cauc	1392.2	405.5	778	2305	
Serum CS846 (ng/mL)	All	95.2	42.0	39	197	1.1	All	95.2	42.0	39	197	1.5
	Female	90.6	39.5	38	183		AA	94.4	31.3	46	174	
	Male	103.2	45.3	41	219		Cauc	95.6	46.7	36	208	
Serum CTXI (ng/mL)	All	0.305	0.224	0.079	0.808	1.4	All	0.305	0.224	0.079	0.808	**1.8**
	Female	0.324	0.245	0.072	0.941		AA	0.375	0.298	0.089	1.031	
	Male	0.271	0.178	0.096	0.598		Cauc	0.268	0.164	0.077	0.685	
Serum COMP (ng/mL)	All	1016.1	650.2	404	2070	1.5	All	1016.1	650.2	404	2070	**2.2**
	Female	1006.1	728.8	381	2091		AA	1114.3	946.6	365	2470	
	Male	1033.6	490.5	447	2031		Cauc	965.3	421.7	428	1876	
Serum HA (ng/mL)	All	36.1	29.4	8.0	104	**1.6**	All	36.1	29.4	8	104	1.2
	Female	32.4	22.8	7.0	94		AA	33.3	26.3	8	92	
	Male	42.6	37.7	8.1	119		Cauc	37.6	30.9	8	111	
Serum MMP-3 (ng/mL)	All	17.0	16.1	3.72	47.89	**2.8**	All	17.0	16.1	3.72	47.89	1.5
	Female	11.1	7.6	3.71	25.39		AA	17.3	11.7	4.38	47.15	
	Male	27.2	21.1	8.53	63.50		Cauc	16.8	18.0	3.43	48.19	
Serum NTXI (nM BCE)	All	16.0	5.1	9	27	1.2	All	16.0	5.1	9	27	1.5
	Female	16.1	4.8	9	27		AA	16.5	6.3	8	30	
	Male	15.9	5.6	8	281.2		Cauc	15.7	4.3	9	26	
Serum PIIANP (ng/mL)	All	2879.9	705.6	1690	4431	1.0	All	2879.9	705.6	1690	4431	1.0
	Female	2904.9	699.8	1750	4390		AA	2811.1	727.7	1642	4338	
	Male	2836.2	721.0	1587	4506		Cauc	2915.5	695.6	1714	4483	
Urine Col2-1 NO2 (nmol/mmol Cr)	All	0.027	0.023	0.0	0.081	**2.5**	All	0.027	0.023	0.0	0.080	**1.7**
	Female	0.031	0.027	0.0	0.097		AA	0.022	0.018	0.0	0.062	
	Male	0.021	0.011	0.002	0.051		Cauc	0.030	0.026	0.0	0.040	
Urine C1,2C (ng/mmol Cr)	All	0.0	0.011	0.010	0.04	1.1	All	0.011	0.010	0.0	0.04	1.2
	Female	0.0	0.011	0.010	0.04		AA	0.014	0.011	0.0	0.05	
	Male	0.0	0.012	0.010	0.05		Cauc	0.010	0.009	0.0	0.03	
Urine C2C (HUSA) (ng/mmol Cr)	All	100.2	67.8	6	268	1.2	All	100.2	67.8	6	268	1.1
	Female	104.4	76.1	4	270		AA	105.9	72.9	7	279	
	Male	97.8	62.9	10	263		Cauc	97.3	65.3	5	262	
Urinary CTXII (ng/mmol Cr)	All	264.3	520.8	50.8	691.3	**6.9**	All	264.3	520.8	50.46	691.34	**4.3**
	Female	188.7	93.9	45.9	857.9		AA	344.2	849.5	47.86	872.45	
	Male	307.6	646.8	64.3	439.1		Cauc	222.9	197.0	52.25	608.43	
Urine NTXI (nmol BCE/mmol Cr)	All	36.9	21.4	11.0	90	1.2	All	36.9	21.4	11	90	1.1
	Female	29.1	17.9	13.0	101		AA	37.6	22.9	10	98	
	Male	41.4	22.0	10.0	64		Cauc	36.5	20.7	12	86	
Urine (α) CTXI (μg/mmol Cr)	All	0.573	0.551	0.04	1.81	1.2	All	0.573	0.551	0.04	1.81	1.4
	Female	0.626	0.507	0.05	2.02		AA	0.640	0.680	0.05	1.98	
	Male	0.479	0.616	0.04	1.44		Cauc	0.538	0.430	0.04	1.73	

Continued

Table 3 Continued

Biomarker	Gender	Mean	SD	Central 95% RI Lower limit	Central 95% RI Upper limit	SD ratio	Race	Mean	SD	Central 95% RI Lower limit	Central 95% RI Upper limit	SD ratio
Urine (β) CTXI (μg/mmol Cr)	All	2.51	1.88	0.46	8.20	1.4	All	2.51	1.88	0.46	8.20	1.1
	Female	2.81	2.03	0.45	9.65		AA	2.69	1.98	0.47	9.10	
	Male	1.97	1.47	0.49	5.74		Cauc	2.41	1.84	0.45	7.81	

SD ratios >1.5 are shown in bold typeface.
Based on 82 females, 47 males and 44 African-Americans (AA), 85 Caucasians (Cauc). Partitioning used the transformed parametric method to compute RIs with the exception of uC1,2C that used the non-parametric method.
BCE, bone collagen equivalents; C1,2C, C-terminal neoepitope of 3/4 piece generated by cleavage of types I and II collagen by collagenases; C2C, C-terminal neoepitope of 3/4 piece generated by cleavage of type II collagen by collagenases; Col2-1 NO2, nitrated type II collagen degradation fragment; COMP, cartilage oligomeric matrix protein; CPII, type II collagen carboxy propeptide cleaved following release of newly synthesised procollagen into matrix; Cr, creatinine; CS846, chondroitin sulfate epitope of aggrecan; CTXI, C-terminal telopeptide of type I collagen (also called CrossLaps), α CTXI contains the newly synthesised α form of aspartic acid, β CTXI contains the age-related β form of aspartic acid; CTXII, C-terminal telopeptide of type II collagen (also called CartiLaps); HA, hyaluronan; MMP-3, matrix metalloproteinase-3; NTXI, N-terminal telopeptide of type I collagen; PIIANP, N-propeptide encoded by exon 2 of type II collagen; RI, reference interval.

OA, the lack of ascertainment of OA status leads to misclassification of controls. These data provide a context for interpretation of OA subject biomarker data and a reference for future human studies. To our knowledge, no comparable 'multijoint control' sample has been characterised previously for such an extensive panel of OA-related biomarkers.

Author affiliations
[1]Division of Rheumatology, Department of Medicine, Duke Molecular Physiology Institute, Duke University School of Medicine, Durham, North Carolina, USA
[2]LabCorp Clinical Trials, San Leandro, California, USA
[3]Rheumatology Department, Royal North Shore Hospital and Institute of Bone and Joint Research, Kolling Institute, University of Sydney, Sydney, New South Wales, Australia
[4]Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[5]Department of Medicine, Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Twitter Follow David Hunter at @ProfDavidHunter

Acknowledgements Scientific and financial support for the FNIH OA Biomarkers Consortium and the study were made possible through grants and direct contributions provided by AbbVie, Amgen Inc., Arthritis Foundation, Bioiberica S.A., DePuy Mitek, Inc., Flexion Therapeutics, Inc., GlaxoSmithKline, Merck Serono, Rottapharm/Madaus, Sanofi and Stryker. In-kind donations to support the biochemical testing were being provided by Alere Inc., ARTIALIS S.A., BioVendor—Laboratorni medicina a.s., IBEX Pharmaceuticals Inc., Immunodiagnostic Systems Ltd and Quidel Corporation. Special thanks to the following: Steve Hoffmann (FNIH), the Scientific Program Manager of the FNIH OA Biomarkers Consortium; W Patrick Gale (UNC Chapel Hill) for data extraction and pulling of samples; Xiaoyan A Shi (UNC Chapel Hill) for identifying the multijoint control sample set; Li Cao and Des Delute (LabCorp Clinical Trials) for performing all the biomarkers assays.

Contributors VBK, JMJ, DEH and DJH designed the study. DEH performed all biomarker analyses. JBR performed all radiographic grading. All authors helped to draft the manuscript and reviewed and approved it for submission.

Funding The Johnston County Osteoarthritis Project has been in part from the Centers for Disease Control and Prevention (CDC) and the Association of Schools of Public Health (ASPH) S043 and S3486 and the National Institute of Arthritis Musculoskeletal and Skin Diseases (NIAMS) Multidisciplinary Clinical Research Center (MCRC) P60 AR49465.

Competing interests DEH is an employee of LabCorp Clinical Trials and was blinded to all clinical data during performance of biomarker assays.

Ethics approval UNC Chapel Hill IRB.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The biomarker data reported in this study are freely available on the Osteoarthritis Initiative (OAI) website (https://oai.epi-ucsf.org/datarelease/FNIH.asp).

REFERENCES

1 Loeser RF, Goldring SR, Scanzello CR, *et al*. Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum* 2012;64:1697–707.
2 Arden N, Nevitt MC. Osteoarthritis: epidemiology. *Best Pract Res Clin Rheumatol* 2006;20:3–25.
3 Glyn-Jones S, Palmer AJ, Agricola R, *et al*. Osteoarthritis. *Lancet* 2015;386:376–87.
4 Kraus V, Collins J, Hargrove D, *et al*. Predictive validity of biochemical biomarkers in knee osteoarthritis - data from the FNIH OA Biomarkers Consortium. *Ann Rheum Dis* 2017;76:186–95.
5 Jordan JM, Helmick CG, Renner JB, *et al*. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. *J Rheumatol* 2007;34:172–80.
6 Kraus VB, Jordan JM, Doherty M, *et al*. The Genetics of Generalized Osteoarthritis (GOGO) study: study design and evaluation of osteoarthritis phenotypes. *Osteoarthritis Cartilage* 2007;15:120–7.
7 Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement* 1977;1:385–401.
8 Eaton W, Muntaner C, Smith C, *et al*. Center for epidemiologic studies depression scale: review and revision (CESD and CESD-R). In: Maruish M, ed. *The use of psychological testing for treatment planning and outcomes assessment*. 3rd edn. Mahwah, NJ: Lawrence Erlbaum, 2004:363–77.
9 Nelson AE, DeVellis RF, Renner JB, *et al*. Quantification of the whole-body burden of radiographic osteoarthritis using factor analysis. *Arthritis Res Ther* 2011;13:R176.
10 Jordan JM, Linder GF, Renner JB, *et al*. The impact of arthritis in rural populations. *Arthritis Care and Res* 1995;8:242–50.
11 Kraus VB, Burnett B, Coindreau J, *et al*. Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. *Osteoarthritis Cartilage* 2011;19:515–42.
12 Elliott AL, Kraus VB, Luta G, *et al*. Serum hyaluronan levels and radiographic knee and hip osteoarthritis in African Americans and Caucasians in the Johnston County Osteoarthritis Project. *Arthritis Rheum* 2005;52:105–11.
13 Rousseau JC, Zhu Y, Miossec P, *et al*. Serum levels of type IIA procollagen amino terminal propeptide (PIIANP) are decreased in patients with knee osteoarthritis and rheumatoid arthritis. *Osteoarthritis Cartilage* 2004;12:440–7.
14 Horowitz G. Defining, establishing, and verifying reference intervals in the clinical laboratory. 3rd edn. Clinical and Laboratory Standards Institute, 2010.