

1 Osteosarcoma: novel prognostic biomarkers using circulating and cell-free tumour DNA

2

3 Iben Lyskjær^{1,2}, Neesha Kara², Solange De Noon^{1,3}, Christopher Davies^{1,3}, Ana Maia Rocha^{1,3},
4 Anna-Christina Strobl³, Inga Usher¹, Craig Gerrand⁴, Sandra J Strauss⁵, Daniel Schrimpf⁶,
5 Andreas von Deimling⁶, Stephan Beck^{2*} & Adrienne M Flanagan^{1,3*#}

6

7 1: Research Department of Pathology, University College London, London, UK

8 2: Medical Genomics Research Group, University College London, London, UK

9 3: Department of Histopathology, Royal National Orthopaedic Hospital, Stanmore, HA7 4LP
10 UK

11 4: Bone Tumour Unit, Royal National Orthopaedic Hospital, Stanmore, HA7 4LP, UK

12 5: Department of Oncology, University College London, London, UK

13 6: Department of Neuropathology, Institute of Pathology, University Hospital Heidelberg, and
14 CCU Neuropathology, German Cancer Institute, Heidelberg, Germany

15

16 *These authors contributed equally

17 #Corresponding author

18

19 Keywords: osteosarcoma, DNA methylation, epigenetics, epigenetic biomarkers, liquid
20 biopsy, prognosis, circulating tumour DNA

21

22 Abstract

23 Osteosarcoma (OS) is the most common primary bone tumour in children and adolescents.
24 Despite treatment with curative-intent, many patients die of this disease. Biomarkers for
25 assessment of disease burden and prognoses for osteosarcoma are not available. Circulating-
26 free (cfDNA) and -tumour DNA (ctDNA) are promising biomarkers for disease surveillance in
27 several major cancer types, however only two such studies are reported for OS. In this
28 combined discovery and validation study, we identified four novel methylation-based
29 biomarkers in 171 OS tumours (test set) and comprehensively validated our findings *in silico*
30 in two independent osteosarcoma sample datasets (n= 162, n=107) and experimentally using
31 digital droplet PCR (ddPCR, n=20 OS tumours). Custom ddPCR assays for these biomarkers
32 were able to detect ctDNA in 40% of pre-operative plasma samples (n=72). ctDNA was
33 detected in 5/17 (29%) post-operative plasma samples from patients who experienced a
34 subsequent relapse post-operatively. Both cfDNA levels and ctDNA detection independently
35 correlated with overall survival, p=0.0015, p=0.0096, respectively. Combining both assays
36 increased the prognostic value of the data. Our findings illustrate the utility of mutation-
37 independent methylation-based markers, broadly applicable ctDNA assays for tumour
38 surveillance and prognostication. This study lays the foundation for multi-institutional
39 collaborative studies to explore the utility of plasma-derived biomarkers for predicting clinical
40 outcome of OS.

41

42

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

43 **1. Introduction**

44 Osteosarcoma (OS) is the most common primary malignant bone tumour in childhood and
45 adolescents, generally arising in the metaphysis of long bones¹. Approximately 80-90% of all
46 OS tumours are high grade, and despite toxic multimodal therapies, survival for the majority
47 of patients with OS has not improved over the last 40 years², with 30-40% of patients with OS
48 dying of their disease³. Although, response to chemotherapy is an indicator of survival⁴, it is
49 currently not possible to predict which patients are likely to respond to chemotherapy.

50

51 Studies have demonstrated the value of minimally invasive surveillance assays in cancer
52 management including assessment of the efficacy of treatment through detection of residual
53 tumour following surgery, and predicting response to chemotherapy or disease relapse⁵.
54 Measurement of total circulating free and tumour DNA (cfDNA and ctDNA, respectively),
55 representing DNA shed into the bloodstream by tumour cells, has been shown to correlate
56 with important clinical endpoints (*e.g.* progression-free survival and or overall survival) in
57 several cancer types, as reviewed in⁶⁻⁸. Barris *et al.* published the first ctDNA study in OS
58 patients (10 patients, 28 samples) using targeted sequencing of genetic alterations detected
59 in the tumours⁹, and found that ctDNA was detected in approximately half of the plasma
60 samples (46%)⁹. Shulman *et al.* used ultra-low-pass whole-genome sequencing (WGS) and
61 detected ctDNA in 57% (41/72 OS patients) of localised OS prior to treatment, and reported
62 that detection of ctDNA was associated with inferior survival but the results were not
63 statistically significant¹⁰.

64

65 In the human genome, DNA methylation predominantly occurs at CpG dinucleotides, the site
66 where a cytosine residue precedes a guanine residue in the 5'-3' direction. In the healthy

67 human genome the majority of CpGs are methylated except at CpG islands¹¹. Conversely, a
68 cancer genome is often characterised by global hypomethylation and focal hypermethylation
69 at CpG islands, leading to both genomic instability and transcriptional repression¹². This
70 feature of cancer has been exploited by studies, which demonstrate ctDNA methylation
71 profiles as a highly sensitive and specific means of detecting cancers¹³. To date, only a few
72 studies have investigated methylation of OS^{14–16}. The use of DNA methylation as a marker of
73 ctDNA in OS is attractive because this tumour is characterised by complex genomes and high
74 inter-tumour heterogeneity rather than recurrent genetic alterations^{17–19}, making selection
75 of a universal genetic ctDNA marker challenging. Furthermore, as changes in methylation
76 represent an early event in cancer and are considered to be retained through tumour
77 evolution²⁰, they could be exploited as powerful biomarkers and would be expected to be
78 only moderately confounded by tumour genetic heterogeneity²¹. For this reason, this
79 approach potentially has advantages over the two previous ctDNA studies in patients with
80 osteosarcoma in which genomic alterations were targeted^{9,10}.

81

82 It has been previously established that high levels of DNA in the circulation (cfDNA) have been
83 linked to inferior outcome^{22–24}. However, in contrast to ctDNA which represents a tumour-
84 specific assay, cfDNA levels can be influenced by other factors including physical activity, age
85 and rate of clearance. For these reasons cfDNA is considered a less reliable biomarker,
86 particularly when measured shortly after surgery⁶. Despite this, the clinical utility of cfDNA as
87 a prognostic marker has been proven in a wide range of cancers⁶.

88

89 The aim of this study was to identify plasma-derived biomarkers in patients with OS that could
90 be used for predicting outcomes.

91 **2. Materials and methods**

92 **2.1. Ethical Approval**

93 The Royal National Orthopaedic Hospital (RNOH) Biobank is approved by the National
94 Research Ethics Committee of the Health Research Committee (reference: Integrated
95 Research Application System (IRAS) project identifier: 272816). This study was approved by
96 the National Research Ethics Committee approved UCL/UCLH Biobank Ethics Committee
97 (project no: EC17.14).

98

99 **2.2. Training methylation cohort**

100 Our *in-house* dataset consisted of 750 samples of various sarcoma subtypes (Supplementary
101 Table 1) (EMBL-EBI, under accession number E-MTAB-9875) previously published²⁵. Infinium
102 450K methylation data was downloaded from TCGA (<https://www.cancer.gov/tcg>) using
103 TCGAbiolinks²⁶ and from Marmal-aid²⁷ (kindly provided by Dr Robert Lowe, UCL, QMUL, UK)
104 (Supplementary Table 2).

105

106 **2.3. External validation cohorts for identified methylation markers**

107 Validation Set I consisted of 450K and EPIC methylation raw IDAT files from 162 OS from
108 Heidelberg, Germany (provided by DS and AVD²⁸, GSE140686) and 699 peripheral blood
109 leucocytes (PBLs) from non-cancer patients (GSE125105²⁹). Validation Set II consisted of 21
110 publicly available OS samples (GSE58770³⁰), 86 OS samples from the TARGET-OS (Children's
111 Oncology Group and The Hospital for Sick Children in Toronto, Canada, dbGaP accession
112 phs000218.v21.p7, <http://ocg.cancer.gov/programs/target> and 732 healthy blood samples
113 (GSE87571³¹).

114

115 **2.4. Biomarker discovery and digital droplet PCR assay design**

116 To avoid false positive results in the plasma samples, we first excluded all CpG sites that
117 showed signs of DNA methylation (β -value > 0.2) in more than 1% of all peripheral blood
118 leucocytes (PBLs) samples (TCGA/Marmal-aid cohorts). Next, we excluded a) CpGs that were
119 hypomethylated (β < 0.5) in more than 20% of all OS cases, b) hypermethylated CpGs (β > 0.5)
120 that were detected in more than 50% of other cancers and c) hypermethylated CpGs (β > 0.2)
121 detected in more than >30% of healthy tissues (Supplementary Figure 1B).

122

123 The 17 identified candidate methylation markers (CpGs) were annotated using the *COHCAP*³²
124 R package (using the hg19 genome build). Methylation-specific digital droplet PCR (ddPCR)
125 assays were designed using Primer3Plus³³ and MethPrimer³⁴, and checked using BiSearch³⁵.
126 Receiver operating characteristics (ROC) analyses were performed using the pROC R
127 package³⁶ to determine the sensitivity and specificity of the probes both individually and in
128 combination (where at least two of the markers were positive). A β -value of 0.5 was chosen
129 as the threshold for determining if a sample should be classified as positive. To be able to use
130 a droplet digital PCR assay with the available plasma (1-2 mL), we reduced the number of
131 markers from 17 to four. This was achieved by first calculating the sensitivities and
132 specificities in the Training and Validation sets (Supplementary Table 8) and then identifying
133 regions for which ddPCR assays could be designed. This led us to selecting the following CpGs
134 for experimental validation; cg02169391, cg22082800, cg25680486, and cg26100986.

135

136 **2.5. Clinical samples for experimental validation**

137 Three clinical sample sets were available for experimental validation: 1) high grade OS tumour
138 samples from 20 patients (Supplementary Table 3), 2) control plasma samples obtained from

139 47 healthy volunteers (2.5 mL plasma) and 69 individuals with benign bone disease (see below
140 for list) and degenerative joint disease (referred to hereafter as non-cancer patients; 4.5 mL
141 plasma, Supplementary Table 4), and 3) samples from 72 patients with OS; these included
142 blood for germline DNA and their pre-operative, pre-treatment plasma samples, in addition
143 to 17 post-OP plasma samples (a total of 89 plasma samples with a median of 4.5 mL, range
144 3-4.5 mL (1-2 mL were used for ddPCR analysis), overview in Supplementary Table 6). The
145 samples of benign bone disease mentioned above included simple bone cysts, osteoarthritis,
146 osteoblastoma, fibrous dysplasia, osteoid osteoma, osteochondroma, tenosynovial giant cell
147 tumour, synovial chondromatosis; the degenerative joint disease represented patients
148 undergoing hip or knee replacements (Supplementary Table 7).

149

150 Clinical correlates

151 Osteosarcoma patients were identified from a research database of patients diagnosed and
152 managed within the London Sarcoma Service and included patients of all ages with
153 demographic data incorporating primary site of disease, tumour grade and size, treatment
154 parameters (chemotherapy, surgery and radiotherapy) as well as pathological response to
155 chemotherapy and outcome (Supplementary Table 5).

156

157 **2.6. Statistical analysis**

158 All statistical analysis was performed in R (version 3.6.1)³⁷. A p-value <5% was deemed
159 statistically significant. A ctDNA-positive sample was defined as a sample with a least two of
160 the four methylation markers above the limit of detection (LOD). Survival analysis was
161 performed using the Kaplan-Meier estimator with death as the end point. The multivariate

162 analysiis was a Cox proportional hazards regression model with time from diagnosis to
163 death or censorship (time to last follow-up) as an end point.

164

165 **2.7. Data availability**

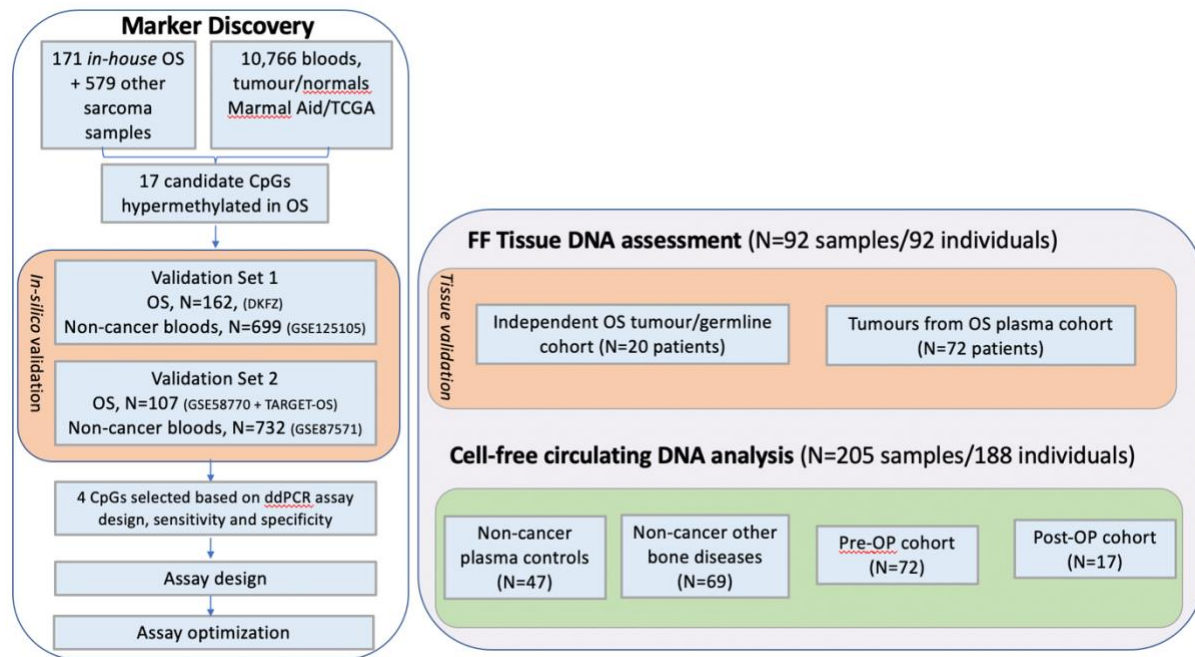
166 *In-house* raw methylation array data have previously been published²⁵ and are available from
167 the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress), accession number E-
168 MTAB-9875.

169

170 **3. Results**

171 **3.1. *In silico* identification and validation of methylation markers through large scale** 172 **bioinformatic analyses**

173 Our first step in the identification of novel CpG sites as candidates for ctDNA assays involved
174 downloading and interrogating the publicly available methylation data from 8,730 and 2,885
175 human tissue samples from the TCGA³⁸ and Marmal-Aid²⁷, respectively (Supplementary Table
176 2, and Figure 1 for study outline). 10,766 samples were eligible for further analysis after
177 filtering for incomplete data (Methods, Supplementary Figure 1A), and were analysed
178 alongside 750 *in-house* sarcoma methylation profiles (Supplementary Figure 1A). Data from
179 the remaining 10,766 samples were used for bioinformatic analysis together with 750 of our
180 *in-house* sarcoma samples²⁵. Hence, the final discovery group included methylome data from
181 11,516 samples, including 1,578 blood samples, 202 normal tissues, 9,565 non-OS cancer
182 samples and 171 OS tumour samples (Supplementary Figure 1, Supplementary Tables 1 and
183 2).



184

185 **Figure 1: Overview of the combined biomarker and validation study.** 17 osteosarcoma-

186 specific candidate methylation markers were identified through a stepwise approach for

187 selecting CpG sites uniquely methylated in OS (see Methods). Four markers were chosen for

188 methylation-specific digital droplet PCR (ddPCR) analysis. The markers were subsequently

189 validated both *in silico* and experimentally in tissue, before being used for circulating tumour

190 DNA assessment of plasma samples from osteosarcoma patients. ‘Other sarcoma’ denotes

191 various other non-OS soft tissue and bone sarcoma subtypes. FF: fresh frozen tumour tissue.

192

193 3.1.1. Osteosarcoma-specific DNA methylation markers

194 We next took a stepwise approach to identify DNA methylation markers (CpG sites), which

195 were uniquely methylated in OS (Supplementary Figure 1B). Using our screening pipeline

196 (Supplementary Figure 1B), we identified 17 candidate hypermethylated markers for OS

197 (Supplementary Figure 2). To obtain an estimate of their combined sensitivity and specificity,

198 we performed receiver operating characteristic (ROC) curve analysis on 171 OS and 1,578

199 non-cancer PBLs (Training Set). The mean area under the curve (AUC) was 0.994 and included
200 all 171 OS and only 11/1578 (0.7%) non-cancer PBLs samples (Supplementary Figure 3A).

201

202 **3.1.2. Candidate DNA methylation markers show robust performance in independent** 203 **OS datasets**

204 Two independent rounds of *in silico* validation of these candidate methylation markers were
205 performed using Validation Set 1 comprising 162 OS tumour and 699 non-cancer PBLs, and
206 Validation Set II consisting of 107 OS tumour and 732 PBL samples (Materials and Methods).
207 By including all 17 candidate markers with a β -value cut-off ≥ 0.5 , all OS in both validation
208 sample sets were correctly classified. None of the 699 PBLs were hypermethylated ($\beta \geq 0.5$)
209 at any of the 17 candidate markers in Validation Set 1, and only two of 732 PBLs (0.27%) from
210 non-cancer patients were misclassified as OS in Validation Set 2. ROC analyses gave an AUC =
211 1.00 for Validation Set 1 and AUC = 0.99 for Validation Set 2 (Supplementary Figure 3B and
212 3C).

213

214 **3.1.3. Four candidate DNA methylation markers were selected for experimental** 215 **validation**

216 Digital droplet PCR-based assays were chosen as the method for ctDNA measurement as this
217 represents a widely employed and highly sensitive, specific, and cost-effective methodology.
218 The amount of plasma available from patients and the number of markers that can be multi-
219 plexed represent limitations of this approach and therefore from the 17 candidate markers,
220 we selected four CpGs (cg02169391, cg22082800, cg25680486, and cg26100986) to take
221 forward (Methods). Hypermethylation of all four markers correctly identified 421/440 of all
222 OS samples studied (95.7%) and misclassified only 6/3009 (false positives = 0.2%) non-cancer

223 PBLs as OS *in silico* (Training and the two Validation Sets) (Supplementary Table 8,
224 Supplementary Figure 4).

225

226 3.2. Experimental validation of methylation markers in clinical OS cohorts

227 3.2.1. Candidate markers are highly sensitive and specific

228 To validate further the sensitivity and specificity of the four markers, we tested these as two
229 duplex ddPCR assays experimentally on samples from another set of 20 high-grade OS and
230 matching germline blood DNA (Supplementary Table 3). Hypermethylation at these CpG sites
231 was confirmed in each of the tumours and in none of the germline samples (Supplementary
232 Figure 5).

233

234 To refine the limit of detection (LOD, Materials and Methods and Supplementary Methods),
235 we also applied the four assays (two duplexed assays) to 47 control plasma samples from
236 patients with non-neoplastic conditions (Supplementary Table 6, Supplementary Figure 5)
237 and plasma from 69 individuals with benign bone disease (Methods, Supplementary Table 7).
238 3/116 samples gave a false positive result (cases 57920 and 59831 (osteochondromas) and
239 63637 (healthy volunteer), with an overall specificity of 97.4%.

240 3.2.2. Circulating tumour DNA detection correlates with disease progression

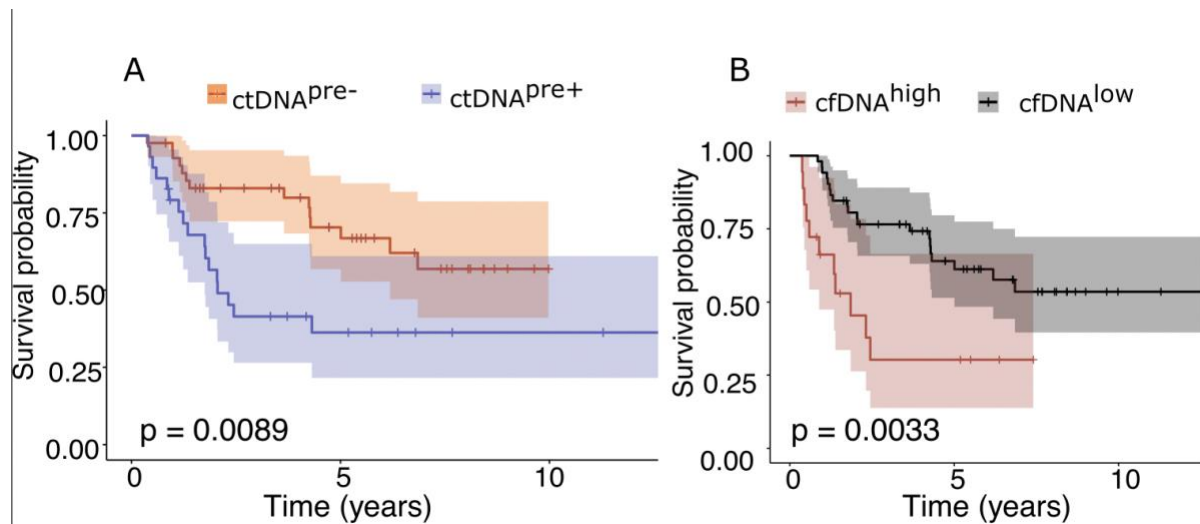
241 Finally, we measured the four methylation markers in matched tumour, PBLs and pre-
242 treatment plasma samples from 72 patients with OS (Supplementary Table 4 and 5). Based
243 on the probes with the highest methylation levels observed in the tumour tissue, we ran one
244 duplexed assay on the matched plasma sample. If neither or only one of the methylation
245 markers was detected, we ran the other duplex assay. Although we observed methylation of

246 at least one marker in 50/72 (69%) samples, we set a minimum of at least two markers to be
247 detected above the LOD for a sample to be classified as 'ctDNA positive' and employed this
248 criterion hereafter.

249

250 Clinical outcome data was then assessed against our pre-operative ctDNA assay results (Table
251 1). Multivariate analysis confirmed that the presence of metastases at diagnosis is an
252 independent prognostic factor in our cohort (Table 2), a finding consistent with the
253 literature³⁹. ctDNA assays were positive in 29/72 (40%) pre-operative samples (ctDNA^{pre+},
254 Table 1, Supplementary table 4). Within this ctDNA^{pre+} group, 12/29 patients had metastatic
255 disease at diagnosis. Of the 17 ctDNA^{pre+} patients with localised disease at diagnosis, 10
256 suffered a disease related event (death, n= 8; or metastasis, n=2). The remaining seven
257 ctDNA^{pre+} patients were alive without disease at last follow up (median follow-up = 5.3 years,
258 range 0.9-11.3 years, Supplementary Table 5). Overall, the detection of pre-treatment ctDNA
259 correlated with inferior survival outcomes (p<0.01, Figure 2A, Table1) and with a disease-
260 related event (p=0.01, chi-square test), but not with tumour size (p=0.14, Student's t-test),
261 gender (0.53) or age at presentation (p=0.92). ctDNA positivity did not have a significant
262 association with outcome on multivariate analysis, however it should be noted that this
263 model also included metastatic disease status at diagnosis which is known to be linked to
264 increased tumour burden^{40,41}, and thus higher probability of ctDNA detection.

265



266

267 **Figure 2: Detection of ctDNA or high levels of cfDNA pre-operatively correlates with survival**

268 A) Detection of ctDNA and B) cfDNA pre-operatively correlate with overall survival (disease-
269 related mortality or censorship). Survival analysis utilised a standard Cox proportional
270 hazard model. The plots are censored after 12 years.

271

272 We also analysed post-operative plasma available from 17 of the 72 patients with OS in whom
273 ctDNA was detected pre-treatment. ctDNA was detected in 5/17 (29%, median days from
274 surgery 18, range: 7-30 days), 4/5 of whom developed recurrent disease (local recurrence or
275 metastasis) and or died of disease (Supplementary Table 5), the remaining patient (OS_66) is
276 alive and well nine years post-surgery. The ctDNA biomarkers were detected in 2/4 patients
277 before the clinical relapse was identified. Detection of ctDNA post-operatively was not
278 associated with the presence of metastasis at diagnosis ($p=0.45$, Student's t-test). Of the
279 patients for whom post-operative samples were available, four and 12 resection specimens
280 were reported as having a good and poor response to chemotherapy, respectively (one
281 unknown), and all cases were reported as having been fully excised.

282

283 3.2.3. Total circulating free DNA levels correlate with outcome

284 Finally, we measured the total amount of cfDNA in the patients' plasma samples, the median
285 level of which was 1,848 copies/mL plasma (range 231-15,180 copies/mL plasma). This
286 correlated with survival (cutoff=75th centile, cfDNA^{high}, p=0.003, Figure 2B), but not with
287 tumour size (rho=0.63, spearman rank test), metastasis at diagnosis (p=0.18, Kruskal-Wallis
288 rank sum test) or anatomical site (femur versus tibia; p=0.63, Kruskal-Wallis rank sum test).
289 12/19 (63%) of those patients classified as cfDNA^{high} were also ctDNA^{pre+}; of the remaining
290 7/19 cfDNA^{high} cases without ctDNA^{pre+}, two had metastasis at diagnosis and another two
291 relapsed (Supplementary Table 5). Including cfDNA^{high} as a risk factor for adverse outcomes
292 alongside ctDNA^{pre+} and metastases at diagnosis resulted in 41/72 (57%) patients classed as
293 high risk at the time of diagnosis. This three risk factor approach correlated with a poor
294 prognosis for this group (p=0.01, Supplementary Figure 6, Supplementary Table 5).

295

296 4. Discussion

297 The standard of care for patients with osteosarcoma² results in an approximate 70% five-year
298 overall survival, thus despite aggressive multi-drug treatment 30-40% die of their disease. The
299 identification of metastases at diagnosis is the most reliable indication that a patient with
300 osteosarcoma has an unfavourable prognosis. However, other factors also contribute to
301 outcome, including tumour size and site, and response to chemotherapy, site^{42,43}.
302 Nevertheless, even when combining all these clinical variables, clinicians lack sufficient
303 information to stratify patients adequately for treatment and to offer patients with OS a
304 reliable prognosis.

305

306 Here, we report a pipeline for the development and application of DNA methylation markers
307 in OS and demonstrate their utility in detecting ctDNA and show that their detection in plasma
308 correlates with inferior clinical outcome. Importantly, these methylation markers were
309 detected in the pre-treatment ctDNA of 32% of patients who did not have metastases at
310 presentation, indicating that this assay has the potential to complement and add value to
311 current criteria for guiding treatment and providing more accurate prognoses to patients.
312 However, 24% of patients with a positive pre-operative ctDNA test remain disease-free, which
313 may in some cases be linked to a follow-up period of less than five years. The cfDNA results
314 confirmed the ctDNA findings in most cases with 63% of patients with a high level of cfDNA
315 also being ctDNA^{pre+}.

316

317 The four selected methylation markers appear to be highly specific and sensitive for OS, being
318 detected in 421/440 (96%, Discovery and Validated Sets) cases analysed *in silico*, and
319 therefore have the potential to be an applicable test for up to 90% of patients with high grade
320 OS. False positive ctDNA results in patients without OS using our methylation markers were
321 rare (2.6%) and should not lead to misdiagnoses or inappropriate management provided
322 results are interpreted in the context of the full clinical picture including medical imaging
323 typically in the form of a multidisciplinary meeting. In this context, our study represents a
324 first step towards the goal of turning cancer including OS into a chronic but manageable
325 disease⁴⁴.

326

327 The reason for failing to detect ctDNA in 60% of patients, particularly in those presenting with
328 metastases or whose disease progressed, is unexplained. Nevertheless, it is noteworthy that
329 relatively similar results (54%⁹ and 43%¹⁰ not detected) were reported in two other studies,

330 which involved different study designs and which employed different technologies for the
331 detection of ctDNA. The study of Barras *et al.* only included ten patients, and that of Shulman
332 *et al.* only studied patients without metastasis at presentation^{9,10}. Furthermore, our study
333 included patients aged between five and 95 years, whereas the other two studies were
334 conducted in patients younger than 35 years^{9,10}. It is possible that the sensitivity of all assays
335 undertaken in the different studies could be improved if larger blood volumes were available
336 for analysis. However, the similar findings across the three studies suggest that the
337 pathobiology of tumours that shed detectable levels of ctDNA may be different from those
338 where this does not occur. This finding is not comparable to the results from other cancers,
339 where ctDNA is detectable in the vast majority of high grade cases in cancers such as lung⁴⁵,
340 ovary⁴⁶ and bladder⁴⁷. Indeed, ctDNA is even detected in small colorectal cancer lesions and
341 is being considered as a screening test for early detection of this disease⁴⁸. ctDNA is also
342 detected in virtually all high grade chondrosarcomas, another bone cancer, but not in low
343 grade cases⁴⁹.

344

345 It would be interesting to test head-to-head the various technologies which have been
346 employed to assess ctDNA in OS patients, but this is difficult to undertake because each assay
347 requires a considerable volume of blood. Nevertheless, a greater understanding of why ctDNA
348 is only detected in some patients may be revealed by interpreting the findings in the context
349 of other features such as the pathology, the medical imaging, as well as multi-omic tumour
350 profiles.

351

352 The limitations of the study are related to the application of the methylation markers to
353 clinical outcome and are largely attributable to the rarity of osteosarcoma, the incidence of

354 which is 0.27/100,000 per year in England⁵⁰. The consequence of this includes the relatively
355 small number of samples analysed within a heterogenous cohort of patients across all age
356 groups, and a relatively short median follow-up of approximately 3.5 years. The small number
357 of post-operative samples, and the absence of an independent cohort in which our findings
358 could be validated are also limitations. The small numbers of samples that were available for
359 analysis post-treatment was disappointing but the finding that ctDNA is detectable prior to
360 presentation of clinical relapse is encouraging and requires further investigation.

361

362 The study highlights the need for multi-centre collaboration, to enable the prospective
363 systematic collection and sharing of patient samples with annotated clinical data, to recruit
364 sufficient patient numbers for validation of these biomarkers and determination of their
365 clinical utility in osteosarcomas, or any other rare disease. Multi-centre studies are fraught
366 with challenges; they are expensive, time-consuming to establish, and it can be difficult to
367 agree and prioritise research aims and studies. Nevertheless, results generated by different
368 groups can add value and provide an opportunity for cross-validation. To this end, the ICONIC
369 study (Improving Outcomes through collaboration in Osteosarcoma)⁵¹, brought together the
370 osteosarcoma clinical and research communities across the UK and provides the
371 infrastructure to prospectively evaluate clinical and biological questions, such as this, to
372 better stratify patients for treatments and improve outcome. It is currently recruiting newly
373 diagnosed OS patients across the UK⁵².

374

375 **Acknowledgements**

376 Funding for this project was received from the Tom Prince Cancer Trust, The Rosetrees and
377 Stoneygate Trusts (M46-F1) and the Bone Cancer Research Trust (BRCT: 6519). IL is supported

378 by the Lundbeck Foundation (R303-2018-3018) and previously by the Tom Prince Cancer
379 Trust. NK was funded by Imperial College London. SB acknowledges support from the
380 Wellcome Trust (WT-218274/Z/19/Z). The study was also supported by the National Institute
381 for Health Research, UCL Genomics and Genome Engineering, UCL Great Ormond Street
382 Institute of Child Health and Pathology Core facilities, the Cancer Research UK University
383 College London Experimental Cancer Medicine Centre, and the RNOH Research and
384 Development Department. AMR and CD was supported by the Tom Prince Cancer Trust. SJS
385 was supported by the National Institute for Health Research, the University College London
386 Hospitals Biomedical Research Centre.

387 The authors thank Dr Sarah Ø. Jensen and Dr Jacob H. Fredsøe, Aarhus University in Denmark,
388 for inputs to the biomarker discovery, Professor Gert Attard and Dr Anjui Wu for a valuable
389 discussion of the experimental design as well as Dr Robert Lowe for sharing data from the
390 Marmal-Aid database. The results generated in this article are partly based upon data
391 generated by the TCGA research Network (<http://cancergenome.nih.gov>), Marmal-Aid
392 database and the Therapeutically Applicable Research to Generate Effective Treatments
393 (TARGET) initiative managed by the NCI. We acknowledge the RNOH Biobank Team and all
394 healthcare workers who cared for the patients. We give a special thanks to the patients for
395 their donating invaluable samples for research.

396

397 **Statement of author contribution**

398 Conceptualization: IL, SB, AMF; Sample curation: AMF; Data curation: IL, NK, AMR, CD, ACS,
399 IU, DS, AVD; Analysis: IL, NK, SDN; Clinical data: AMF, SDN, CG, SJS; Supervision: SB, AMF;
400 Writing: IL, SB, AMF, SDN; Review and editing: All authors.

401

402 References

- 403 1. Misaghi, A., Goldin, A., Awad, M. & Kulidjian, A. A. Osteosarcoma: a comprehensive
404 review. *SICOT-J* (2018). doi:10.1051/sicotj/2017028
- 405 2. Gerrand, C. *et al.* UK guidelines for the management of bone sarcomas. *Clin. Sarcoma*
406 *Res.* **6**, 7 (2016).
- 407 3. Durfee, R. A., Mohammed, M. & Luu, H. H. Review of Osteosarcoma and Current
408 Management. *Rheumatol. Ther.* (2016). doi:10.1007/s40744-016-0046-y
- 409 4. Kager, L. *et al.* Primary metastatic osteosarcoma: Presentation and outcome of
410 patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols. *J.*
411 *Clin. Oncol.* (2003). doi:10.1200/JCO.2003.08.132
- 412 5. Kilgour, E., Rothwell, D. G., Brady, G. & Dive, C. Liquid Biopsy-Based Biomarkers of
413 Treatment Response and Resistance. *Cancer Cell* (2020).
414 doi:10.1016/j.ccell.2020.03.012
- 415 6. Bronkhorst, A. J., Ungerer, V. & Holdenrieder, S. The emerging role of cell-free DNA as
416 a molecular marker for cancer management. *Biomolecular Detection and*
417 *Quantification* (2019). doi:10.1016/j.bdq.2019.100087
- 418 7. Morris, V. K. & Strickler, J. H. Use of Circulating Cell-Free DNA to Guide Precision
419 Medicine in Patients with Colorectal Cancer. *Annual Review of Medicine* (2021).
420 doi:10.1146/annurev-med-070119-120448
- 421 8. Cescon, D. W., Bratman, S. V., Chan, S. M. & Siu, L. L. Circulating tumor DNA and liquid
422 biopsy in oncology. *Nat. Cancer* (2020). doi:10.1038/s43018-020-0043-5
- 423 9. Barris, D. M. *et al.* Detection of circulating tumor DNA in patients with osteosarcoma.
424 *Oncotarget* (2018). doi:10.18632/oncotarget.24268
- 425 10. Shulman, D. S. *et al.* Detection of circulating tumour DNA is associated with inferior
426 outcomes in Ewing sarcoma and osteosarcoma: a report from the Children's
427 Oncology Group. *Br. J. Cancer* (2018). doi:10.1038/s41416-018-0212-9
- 428 11. Suzuki, M. M. & Bird, A. DNA methylation landscapes: Provocative insights from
429 epigenomics. *Nature Reviews Genetics* (2008). doi:10.1038/nrg2341
- 430 12. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome-biological and
431 translational implications. *Nature Reviews Cancer* (2011). doi:10.1038/nrc3130
- 432 13. Koch, A. *et al.* Analysis of DNA methylation in cancer: Location revisited. *Nature*
433 *Reviews Clinical Oncology* (2018). doi:10.1038/s41571-018-0004-4

- 434 14. Chen, X. G., Ma, L. & Xu, J. X. Abnormal DNA methylation may contribute to the
435 progression of osteosarcoma. *Mol. Med. Rep.* (2018). doi:10.3892/mmr.2017.7869
- 436 15. Tian, W., Li, Y., Zhang, J., Li, J. & Gao, J. Combined analysis of DNA methylation and
437 gene expression profiles of osteosarcoma identified several prognosis signatures.
438 *Gene* (2018). doi:10.1016/j.gene.2018.01.093
- 439 16. Rosenblum, J. M. *et al.* Predictive properties of DNA methylation patterns in primary
440 tumor samples for osteosarcoma relapse status. *Epigenetics* (2015).
441 doi:10.4161/15592294.2014.989084
- 442 17. Behjati, S. *et al.* Recurrent mutation of IGF signalling genes and distinct patterns of
443 genomic rearrangement in osteosarcoma. *Nat. Commun.* (2017).
444 doi:10.1038/ncomms15936
- 445 18. Wu, C. C. *et al.* Immuno-genomic landscape of osteosarcoma. *Nat. Commun.* (2020).
446 doi:10.1038/s41467-020-14646-w
- 447 19. Bousquet, M. *et al.* Whole-exome sequencing in osteosarcoma reveals important
448 heterogeneity of genetic alterations. *Ann. Oncol.* (2016).
449 doi:10.1093/annonc/mdw009
- 450 20. Klutstein, M., Nejman, D., Greenfield, R. & Cedar, H. DNA methylation in cancer and
451 aging. *Cancer Research* (2016). doi:10.1158/0008-5472.CAN-15-3278
- 452 21. Gai, W. & Sun, K. Epigenetic Biomarkers in Cell-Free DNA and Applications in Liquid
453 Biopsy. *Genes (Basel)*. (2019). doi:10.3390/genes10010032
- 454 22. Tissot, C. *et al.* Circulating free DNA concentration is an independent prognostic
455 biomarker in lung cancer. *Eur. Respir. J.* (2015). doi:10.1183/13993003.00676-2015
- 456 23. Spindler, K. L. G., Pallisgaard, N., Vogelius, I. & Jakobsen, A. Quantitative cell-free
457 DNA, KRAS, and BRAF mutations in plasma from patients with metastatic colorectal
458 cancer during treatment with cetuximab and irinotecan. *Clin. Cancer Res.* (2012).
459 doi:10.1158/1078-0432.CCR-11-0564
- 460 24. Lyskjær, I. *et al.* Correlation between early dynamics in circulating tumour DNA and
461 outcome from FOLFIRI treatment in metastatic colorectal cancer. *Sci. Rep.* **9**, (2019).
- 462 25. Lyskjær, I., De Noon, S., Tirabosco, R. DNA methylation-based profiling of bone and
463 soft tissue tumours: a validation study of the 'DKFZ Sarcoma Classifier'. *J. Pathol. Clin.*
464 *Res.* (2021). doi:<https://doi.org/10.1002/cjp2.215>
- 465 26. Colaprico, A. *et al.* TCGAblinks: An R/Bioconductor package for integrative analysis

- 466 of TCGA data. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkv1507
- 467 27. Lowe, R. & Rakyan, V. K. Marmalaid—a database for Infinium HumanMethylation450.
468 *BMC Bioinformatics* (2013). doi:10.1186/1471-2105-14-359
- 469 28. Koelsche, C. *et al.* Sarcoma classification by DNA methylation profiling. *Nat. Commun.*
470 **12**, (2021).
- 471 29. Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets - Update.
472 *Nucleic Acids Res.* (2013). doi:10.1093/nar/gks1193
- 473 30. Moriarity, B. S. *et al.* A Sleeping Beauty forward genetic screen identifies new genes
474 and pathways driving osteosarcoma development and metastasis. *Nat. Genet.* (2015).
475 doi:10.1038/ng.3293
- 476 31. Johansson, Å., Enroth, S. & Gyllenstein, U. Continuous Aging of the Human DNA
477 Methylome Throughout the Human Lifespan. *PLoS One* (2013).
478 doi:10.1371/journal.pone.0067378
- 479 32. Warden, C. D. *et al.* COHCAP: An integrative genomic pipeline for single-nucleotide
480 resolution DNA methylation analysis. *Nucleic Acids Res.* (2013).
481 doi:10.1093/nar/gkt242
- 482 33. Untergasser, A. *et al.* Primer3—new capabilities and interfaces. *Nucleic Acids Res.*
483 (2012). doi:10.1093/nar/gks596
- 484 34. Li, L. C. & Dahiya, R. MethPrimer: Designing primers for methylation PCRs.
485 *Bioinformatics* (2002). doi:10.1093/bioinformatics/18.11.1427
- 486 35. Arányi, T. & Tusnády, G. E. BiSearch: ePCR tool for native or bisulfite-treated genomic
487 template. *Methods Mol. Biol.* (2007). doi:10.1385/1-59745-528-8:385
- 488 36. Robin, X. *et al.* pROC: An open-source package for R and S+ to analyze and compare
489 ROC curves. *BMC Bioinformatics* (2011). doi:10.1186/1471-2105-12-77
- 490 37. R Core Team. R: A language and environment for statistical computing.
491 <http://www.R-project.org/>. *R Foundation for Statistical Computing, Vienna, Austria*
492 (2017).
- 493 38. Tomczak, K., Czerwińska, P. & Wiznerowicz, M. The Cancer Genome Atlas (TCGA): An
494 immeasurable source of knowledge. *Wspolczesna Onkologia* (2015).
495 doi:10.5114/wo.2014.47136
- 496 39. SALAH, S., AHMAD, R., SULTAN, I., YASER, S. & SHEHADEH, A. Osteosarcoma with
497 metastasis at initial diagnosis: Current outcomes and prognostic factors in the context

- 498 of a comprehensive cancer center. *Mol. Clin. Oncol.* (2014).
499 doi:10.3892/mco.2014.325
- 500 40. Iacovelli, R., Lanoy, E., Albiges, L. & Escudier, B. Tumour burden is an independent
501 prognostic factor in metastatic renal cell carcinoma. *BJU Int.* (2012).
502 doi:10.1111/j.1464-410X.2012.11518.x
- 503 41. Gómez, O. H. *et al.* "High Tumor Burden" in Metastatic Non-Small Cell Lung Cancer:
504 Defining the Concept. *Cancer Manag. Res.* (2021). doi:10.2147/CMAR.S302928
- 505 42. Bramer, J. A. M., van Linge, J. H., Grimer, R. J. & Scholten, R. J. P. M. Prognostic factors
506 in localized extremity osteosarcoma: A systematic review. *European Journal of*
507 *Surgical Oncology* (2009). doi:10.1016/j.ejso.2009.01.011
- 508 43. Xin, S. & Wei, G. Prognostic factors in osteosarcoma: A study level meta-analysis and
509 systematic review of current practice. *J. Bone Oncol.* (2020).
510 doi:10.1016/j.jbo.2020.100281
- 511 44. Beck, S. & Ng, T. C2c: Turning cancer into chronic disease. *Genome Medicine* (2014).
512 doi:10.1186/gm555
- 513 45. Wang, H. *et al.* The Role of Circulating Tumor DNA in Advanced Non-Small Cell Lung
514 Cancer Patients Treated With Immune Checkpoint Inhibitors: A Systematic Review
515 and Meta-Analysis. *Frontiers in Oncology* (2021). doi:10.3389/fonc.2021.671874
- 516 46. Asante, D. B., Calapre, L., Ziman, M., Meniawy, T. M. & Gray, E. S. Liquid biopsy in
517 ovarian cancer using circulating tumor DNA and cells: Ready for prime time? *Cancer*
518 *Letters* (2020). doi:10.1016/j.canlet.2019.10.014
- 519 47. Birkenkamp-Demtröder, K. *et al.* Longitudinal assessment of multiplex patient-specific
520 ctDNA biomarkers in bladder cancer for diagnosis, surveillance and recurrence. *Ann.*
521 *Oncol.* (2018). doi:10.1093/annonc/mdy269.084
- 522 48. Reinert, T. *et al.* Analysis of Plasma Cell-Free DNA by Ultradeep Sequencing in
523 Patients with Stages I to III Colorectal Cancer. *JAMA Oncol.* (2019).
524 doi:10.1001/jamaoncol.2019.0528
- 525 49. Lyskjær, I. *et al.* Circulating tumour DNA is a promising biomarker for risk
526 stratification of central chondrosarcoma with IDH1/2 and GNAS mutations. *Mol.*
527 *Oncol.* (2021). doi:10.1002/1878-0261.13102
- 528 50. R., G. *et al.* UK guidelines for the management of bone sarcomas. *Sarcoma* (2010).
529 doi:<http://dx.doi.org/10.1155/2010/317462>

- 530 51. ICONIC: Improving Outcomes Through Collaboration in Osteosarcoma (ICONIC).
531 Available at: <https://clinicaltrials.gov/ct2/show/NCT04132895>.
- 532 52. ICONIC – Improving Outcomes through Collaboration In Osteosarcoma. Available at:
533 <https://www.bcrct.org.uk/research/current-research-projects/iconic>.
534
535
536

537 **Tables**

538 Table 1. Correlation of clinical outcome with pre-treatment ctDNA. PR: poor responder; GR:
539 good responder, NCG: No chemo given, ctDNA: circulating tumour DNA.

	ctDNA pre neg (n=43)	ctDNA pre pos (n=29)
Metastasis at diagnosis (n=19)	7 (16%)	12 (41%)
No metastases at diagnosis (n=53)	36 (68%)	17 (32%)
Die of disease (n=32)	14 (33%)	18 (62%)
Disease-related event (n=39)	18 (42%)	21 (72%)
Response to chemotherapy	PR: 28 (65%) GR: 9 (21%) NCG:6 (14%)	PR: 11 (38%) GR: 8 (28%) NCG:4 (14%) Unknown/NA: 5

540

541 **Table 2:** Uni-and multivariate analysis of clinical factors and their relation to survival. cfDNA:
542 circulating free DNA; ctDNA: circulating tumour DNA; pre-OP: pre-operatively.

SURVIVAL	Univariate analysis			Multivariate analysis		
	HR	95% CI	P-value	HR	95% CI	P-value
Pre-OP cfDNA (n=72)						
<75th centile (n=53)	0.34	0.38	0.005	0.56	0.43	0.18
>75th centile (n=19)	REFERENCE					
Pre-OP ctDNA (n=72)						
Neg (n=43)	REFERENCE					
Pos (n=29)	2.51	0.36	0.01	1.48	0.42	0.36
Metastasis at presentation (n=72)						
Yes (n=19)	5	0.37	1.2*10⁻⁵	3.78	0.40	0.0096
No (n=53)	REFERENCE					
Gender (n=72)						

Male (n=38)	REFERENCE					
Female (n=34)	1.16	0.36	0.68			
Age (n=72)						
<30 years (n=62)	REFERENCE					
=>30 years (n=10)	1.14	0.54	0.81			
Max tumour size (n=64)						
<=80 mm (n=14)	REFERENCE					
>80 mm (n=50)	1.38	0.5	0.51			

543