

Body composition and lung cancer-associated cachexia in TRACERx

Othman Al-Sawaf^{1,2,3#}, Jakob Weiss^{4,5#}, Marcin Skrzypski^{6#}, Jie Min Lam^{1,2,25}, Takahiro Karasaki^{11,2,3}, Francisco Zambrana⁷, Andrew C Kidd⁸, Alexander Frankell^{1,3}, Thomas B. K. Watkins^{1,3}, Carlos Martinez-Ruiz¹, Mateo Sokac^{9,10}, Susie Collins¹¹, Selvaraju Veeriah¹, Neil Magno¹, Cristina Naceur-Lombardelli¹, Paulina Prymas¹, Antonia Toncheva¹, Nick Jayanth¹², Roberto Salgado¹³, Christopher P. Bridge¹⁴, David Christiani¹⁵, Raymond Mak¹⁶, Camden Bay¹⁷, Michael Rosenthal¹⁷, Naveed Sattar¹⁸, Paul Welsh¹⁸, Ying Liu^{19,20}, Norbert Perrimon^{19,20}, Karteek Poporui²¹, Mirza Faisal Beg²², Nicholas McGranahan¹, Allan Hackshaw¹², Danna M Breen²³, Stephen O'Rahilly²⁴, Nicolai Juul Birkbak^{9,10}, Hugo Aerts⁴, TRACERx Consortium, Mariam Jamal-Hanjani^{1,2,25}, Charles Swanton^{1,3,25}

Shared first-authors

* co-corresponding authors

- 1 Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK
- 2 Cancer Metastasis Laboratory, University College London Cancer Institute, London, UK
- 3 Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK
- 4 Artificial Intelligence in Medicine Program, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; Cardiovascular Imaging Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.
- 5 Department of Diagnostic and Interventional Radiology, University Freiburg, Germany
- 6 Medical University of Gdańsk, Department of Oncology and Radiotherapy, 17 Smoluchowskiego St., 80-001 Gdańsk, Poland
- 7 Infanta Sofia University Hospital, Madrid, Spain
- 8 Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow (Glasgow), United Kingdom
- 9 Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
- 10 Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark
- 11 Early Clinical Development, Pfizer UK Ltd, Cambridge, UK
- 12 Cancer Research UK & University College London Cancer Trials Centre, London, UK
- 13 Breast Cancer Translational Research Laboratory, Institut Jules Bordet; Department of Pathology/TCRU GZA Antwerp, Belgium
- 14 MGH & BWH Center for Clinical Data Science, Boston, MA, USA
- 15 Harvard Education and Research Center for Occupational Safety and Health, Boston, MA, USA
- 16 Radiation Oncology Disease Center, Brigham and Women's Hospital, Boston, MA, USA
- 17 Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA
- 18 School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, United Kingdom
- 19 Department of Genetics, Harvard Medical School, Boston, United States
- 20 Howard Hughes Medical Institute, Harvard Medical School, Boston, United States
- 21 Department of Computer Science, Memorial University of Newfoundland, St. John's, NL, Canada
- 22 School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada
- 23 Internal Medicine Research Unit, Pfizer Inc, Cambridge, MA, USA
- 24 Wellcome Trust-MRC Institute of Metabolic Science and NIHR Cambridge Biomedical Research Centre, University of Cambridge, Cambridge, UK
- 25 Department of Medical Oncology, University College London Hospitals, London, UK

ABSTRACT

Cancer-associated cachexia (CAC) is a major determinant of morbidity and mortality in patients with non-small cell lung cancer (NSCLC). Key features of CAC include alterations of body composition and body weight. Here, we explore the association between body composition and body weight with survival and delineate possible biological processes and mediators that contribute to the development of CAC. To this end, we report the body composition profiles at diagnosis and relapse in relation to clinical outcomes for patients prospectively recruited into the TRACERx study and investigate the relationship between changes in body composition indicative of CAC with tumour genomic and transcriptomic profiles, complemented by plasma proteomics. Computed tomography-based (CT) body composition analysis of 650

patients in TRACERx suggested that patients with low skeletal muscle or adipose tissue area at the time of lung cancer diagnosis, represented by the bottom 20th percentile, had significantly shorter lung cancer-specific survival (LCSS) and overall survival (OS). This was validated in 420 patients in the independent Boston Lung Cancer Study, in which patients in the bottom 20th percentile similarly had shorter LCSS and OS. In a longitudinal subset of 272 patients in TRACERx who suffered disease relapse, loss of adipose tissue, skeletal muscle, or body weight in the interval between diagnosis and relapse, was significantly associated with shorter LCSS and OS. Subgroups of patients characterised by either predominant loss of adipose or muscle tissue were identified, suggesting possibly distinct clinical subtypes of CAC. Patients with one or more features encompassing loss of adipose or muscle tissue, or BMI-adjusted weight loss according to specific thresholds were classified as having developed CAC and were found to have distinct tumour genomic and transcriptomic profiles compared with patients who did not develop such features at relapse. Primary NSCLC tumours from patients in the CAC group were characterised by enrichment of inflammatory signalling and epithelial-mesenchymal transitional pathways, and differentially expressed genes upregulated in these tumours included *LBP* and matrix metalloproteinases, such as *ADAMTS3*. In an exploratory analysis of putative circulating cachexia mediators performed in a subset of 256 plasma samples from TRACERx, proteomic analysis revealed a significant association between circulating GDF15 with loss of body weight, skeletal muscle, and adipose tissue at relapse, supporting the potential therapeutic relevance of targeting GDF15 in the management of CAC.

INTRODUCTION

Measures of body composition that distinguish skeletal muscle (SKM), visceral adipose tissue (VAT), and subcutaneous adipose tissue (SAT) are associated with clinical outcomes in various diseases, including cancer {Martin, 2013 #1604;Calle, 2003 #1557;Guo, 2016 #1558;Lavie, 2014 #1559}. One extreme manifestation of altered body composition that remains poorly understood is cancer-associated cachexia (CAC); a paraneoplastic syndrome of involuntary skeletal muscle and adipose tissue loss, accompanied by dysregulation of the homeostatic mechanisms that govern protein and energy balance {Baracos, 2018 #1438;Fearon, 2011 #1444}.

Retrospective analyses have linked body composition to outcomes across multiple solid tumour types, including breast, prostate and colorectal cancers {Caan, 2018 #1563} {Lee, 2018 #1564}; and in non-small cell lung cancer (NSCLC), SKM wasting has been shown to be associated with cancer treatment toxicity and reduced overall survival (OS) {Baracos, 2010 #1575}. A meta-analysis of 13 NSCLC cohort studies demonstrated an association between low SKM mass and reduced OS, but not disease-specific survival{Yang, 2019 #1520}; whilst retrospective studies have similarly suggested an association between low SAT {Popinat, 2019 #1576} and low VAT with reduced OS {Popinat, 2019 #1576;{Tan, 2022 #1574}, again without a significant difference in cancer-specific survival. There remain key areas of uncertainty; not least the dynamics of body composition with time, and the underlying molecular mediators of altered body composition in CAC.

Our study had three primary aims: (1) to systematically profile the body composition of patients diagnosed with early-stage lung cancer in terms of SKM, VAT and SAT and to explore associations with survival outcomes (2) to examine how

these features change in the context of cancer recurrence as an indicator of CAC and (3) to explore the corresponding tumour genomic, transcriptomic, and plasma proteomic landscape for possible molecular mechanisms and mediators of CAC. To this end, we used established computed tomography (CT) imaging analysis methods {Shen, 2004 #1560;Shen, 2004 #1561;Mourtzakis, 2008 #1562} to capture body composition at lung cancer diagnosis and relapse, addressing the relevance of initial body composition metrics, and their perturbation at time of recurrence, to lung cancer outcomes and the cachectic phenotype. To investigate the interplay between tumour biology, body composition losses, and clinical outcomes, our analysis plan involved high-dimensional whole exome and transcriptomic analysis of bulk tumour samples, allied to plasma proteomics; to identify tumour-intrinsic factors, alongside potential circulating mediators, underpinning CAC.

We co-opted CT imaging from two large independent cohorts of patients diagnosed with early-stage operable NSCLC, treated with surgical resection, plus adjuvant therapy if indicated. These are the TRACERx (TRACKing non-small cell lung Cancer Evolution through therapy (Rx)) Lung{Jamal-Hanjani, 2017 #1318} study, forming our principal cohort, and the Boston Lung Cancer Study (BLCS){Christiani, 2022 #1577}, forming our validation cohort. Altogether, these encompassed a combined 839 diagnostic CT scans paired with detailed clinical annotation, including outcomes. SAT, VAT and SKM areas at the third lumbar vertebrae level {Anyene, 2022 #1609} were determined using a deep-learning based imaging analysis pipelines, establishing the baseline body composition database. In 272 patients from TRACERx, matched relapse body composition and/or body weight data were further evaluated to identify patients with changes indicative of CAC; primary bulk tumour samples from initial resection were analysed to infer the corresponding tumour

genomic and transcriptomic landscape pertaining to the CAC phenotype. Finally, we sought to establish the presence of circulating CAC mediators by conducting plasma proteomic analysis on a subset of 128 patients from TRACERx (256 plasma samples collected at diagnosis and first relapse) to determine their differential expression in relation to our observed cachexia phenotype. By integrating longitudinal imaging, tumour, and blood analyses, we provide unique insights into the relationship between body composition and cachexia in NSCLC, characterising the tumour genome, transcriptome, and plasma proteome (Figure 1 A), establishing a platform for downstream validation and potential clinical translation.

RESULTS

Low muscle and adipose tissue areas at diagnosis are associated with shorter survival

The first aim of the study was to profile the body composition, based on SAT, VAT and SKM tissue areas, of patients at the time of early-stage NSCLC diagnosis and to elucidate the prognostic impact of low SAT, VAT and SKM tissue areas on cancer specific-survival as the primary outcome measure. Body composition in 651 patients from the TRACERx study and 420 patients from the BLCS study was assessed at the time of cancer diagnosis. All patients had stage I-III disease and underwent primary resection. In the TRACERx cohort, 35% of the patients received adjuvant therapy (55% in the BLCS cohort) and 14% were current smokers (37% in the BLCS cohort), otherwise the baseline characteristics were similar between the two cohorts (Table 1).

Body composition was measured by quantification of tissue area (cm²) at the level of the third lumbar vertebra (L3) on CT, or CT-PET, scan. SAT, VAT and SKM

areas were quantified separately using deep-learning-based, automated pipelines (see Methods) {Dabiri, 2020 #1510; Bridge, 2018 #1512}. The distribution of body SAT, VAT, SKM and BMI was similar in the TRACERx and BCLS cohorts (Figure 1 B, C). In TRACERx, male patients had higher VAT and SKM areas compared with female patients (mean 171.7 cm² vs 95.1 cm², unpaired two-samples t-test p<0.001 and 147.4 cm² vs 99.5 cm², p<0.001), whereas female patients had higher SAT areas (201.2 cm² vs 144.6 cm², p<0.001) (Figure 1 B). There was a strong correlation between BMI and both SAT and VAT (Spearman's correlation, r = 0.75 and 0.73, respectively), and a weaker correlation with SKM (r = 0.39) (Figure S1 A); similar correlations were observed in the BCLS cohort (Figure S1 B).

In order to identify patients with particularly low, normal, or high values/areas for adipose and muscle tissue, patients in the TRACERx cohort were grouped into sex-adjusted bottom 20th, middle 20-80th and upper 20th percentiles. In the TRACERx cohort, patients in the bottom 20th percentile for SAT, VAT and SKM had a significantly shorter lung cancer-specific survival (LCSS) compared to patients in the middle 20-80th percentile, a proxy for patients with 'normal' body composition to mitigate outliers (Figure 1D). After adjusting for age, sex, BMI, smoking status, disease stage, histological subtype, ethnicity, and adjuvant therapy use, all three body composition measures were associated with shorter LCSS. The adjusted hazard ratios for LCSS for patients in the lowest 20th percentile were 2.09 (95% CI 1.55-3.22, p<0.001), 1.73 (1.10-2.72, p=0.019) and 1.44 (0.95-2.19, p=0.088) for SAT, VAT and SKM (Table S1). As with LCSS, a similar association between body composition and shorter overall survival (OS) was identified in patients in the bottom 20th percentile for SAT, VAT and SKM, compared with patients in the middle 20-80th percentile with adjusted hazard ratios of 1.49 (1.02-2.16, p=0.037), 1.38 (0.95-2.01,

p=0.093) and 1.28 (0.91-1.78, p=0.151) for SAT, VAT and SKM, respectively (Table S2).

These observations were validated in the BLCS cohort, where patients in the bottom 20th percentile for SAT, VAT and SKM had a shorter LCSS than patients in the middle 20-80th percentile, similarly adjusted for age, sex, BMI, smoking status, disease stage, histological subtype, ethnicity, and adjuvant therapy use with hazard ratios of 1.97 (1.24-3.14, p=0.004), 1.57 (0.98-2.53, p=0.06), and 1.35 (0.86-2.12, p=0.19), respectively (Figure 1 E, Table S3). For OS, the corresponding hazard ratios were 1.71 (95% CI 1.18-2.49, p=0.005), 1.75 (95% CI 1.20-2.54, p=0.003), and 1.41 (95% CI 1.00-2.01, p=0.05) (Table S4).

In summary, these analyses indicate that patients with NSCLC who present with low SAT, VAT or SKM tissue at the time of cancer diagnosis have a shorter survival than patients with normal or high adipose and muscle tissue. This highlights the prognostic impact of baseline body composition in the context of early-stage NSCLC.

Longitudinal changes in body composition and body weight correlate with outcomes and identify the cachexia phenotype

Next, changes in body composition from diagnosis to relapse (median time to first recurrence 15.6 months, 95% CI 13.9-18.0) were examined by analysing CT images from 188 patients in TRACERx with a confirmed cancer recurrence and available abdominal CT image sections for body composition profiling at these two timepoints (Figure S10). Absolute (cm²) changes, i.e. gains and losses in SKM, VAT, and SAT were calculated, in addition to BMI-adjusted weight loss {Martin, 2014 #1454} (Figure 2 A, B).

To obtain prognostically relevant thresholds of SAT, VAT and SKM loss with minimum significant change, cut points from 5% to 40% were tested for each compartment (Figure S2). A loss of $\geq 20\%$ tissue area from diagnosis to relapse in SAT or VAT was associated with significantly shorter LCSS (SAT: HR 1.56 [95% CI 1.02-2.38], $p = 0.042$; VAT: HR 2.34 [95% CI 1.38-3.99], $p=0.0017$) and shorter OS (SAT: HR 1.59 [95% CI 1.079-2.36], $p=0.019$; VAT: HR 2.39 [95% CI 1.47-3.89], $p<0.001$) (Figure S3). A loss of $\geq 10\%$ skeletal muscle tissue area from diagnosis to relapse was associated with significantly shorter LCSS and OS (HR 1.80 [95% CI 1.20-2.70], $p=0.0047$ and HR 1.85 [95% CI 1.27-2.69], $p=0.0013$). In addition to changes in body composition, grade 4 BMI-adjusted weight (see Methods) loss was significantly associated with shorter LCSS and OS compared to patients with stable body weight (HR 5.637 [95% CI 3.19-9.95], $p<0.0001$ and HR 4.71 [95% CI 2.76-8.04], $p<0.0001$). A strong correlation was observed between body weight loss and VAT loss, and between VAT and SAT loss (Spearman's $r = 0.70$ and 0.62 , Figure S1 C). Likewise, grade 4 BMI-adjusted weight loss was associated with loss of SAT, VAT and SKM (Figure S1 D, $p<0.0001$).

To explore whether distinct patterns of loss affecting specific body composition compartments were present, patients were grouped according to co-occurrence of SKM, VAT, and SAT losses for the above prognostic thresholds (Figure 2C). Amongst 96 patients in whom a loss occurred in any given compartment, 23/96 (24%) patients experienced losses across all three compartments, 25/96 (26%) patients experienced losses across two compartments and 48/96 (50%) patients experienced isolated loss in one compartment, i.e., isolated loss of SKM (N=21), VAT (N=22) or SAT (N=5), suggesting possibly distinct clinical subtypes of CAC with different anatomical patterns of fat and muscle loss.

Notably, patients with tissue loss across all three compartments had the shortest LCSS (HR: 2.6, 95% CI 1.39-4.87, $p=0.003$) (Figure 2C, Table S6).

Given that loss of muscle and adipose tissue are considered hallmark features of CAC, we leveraged the above prognostic thresholds for change in body composition and weight loss to define a group of patients in TRACERx with the CAC phenotype, characterised by $\geq 20\%$ SAT and/or VAT loss, and/or $\geq 10\%$ SKM loss, and/or grade 4 BMI-adjusted weight loss in the interval between diagnosis and relapse. Based on this stratification of patients into CAC and non-CAC groups, a higher proportion of male patients and squamous cell carcinoma histology subtype was observed in the CAC group compared to the non-CAC group (62.0% versus 54.2%, and 34.3% versus 25.8%), whereas smoking status and use of adjuvant therapy was similar between the two groups (Table S5). Patients in the CAC group had a significantly shorter LCSS than patients in the non-CAC group (Figure 2D). Time to recurrence was significantly shorter in patients in the CAC compared with the non-CAC group (median 13.3 months [95% CI 8.9-16.1] vs 18.4 months [15.2-21.0]; HR 1.81, 95% CI 1.38-2.36, $p<0.001$) (Figure S4). This assignment of CAC was an independent prognostic factor for both LCSS and OS when adjusted for sex, ethnicity, BMI, smoking status, histology, disease stage, and use of adjuvant therapy (HR 2.42 [95% CI 1.69-3.46], $p<0.001$ and HR 2.31 [95% CI 1.66- 3.20], $p<0.001$) (Figure 2E, Table S7).

These observations suggest that losses of adipose and muscle tissue as well as high grade BMI-adjusted weight loss are associated with a shorter survival in patients with NSCLC. A combined loss of SAT, VAT and SKM was particularly indicative of a poor prognosis. Integrating significant changes of body composition and body weight identified patients experiencing CAC between primary diagnosis

and first relapse. This annotation provided an opportunity to conduct further downstream analyses to develop biological hypotheses on possible drivers of CAC.

Primary tumour genomic and transcriptomic features in relation to body composition and body weight changes at first cancer relapse

In order to examine the presence of genomic and transcriptomic alterations potentially relevant to the development of CAC, we analysed tumour sequencing data in relation to body composition changes between diagnosis and relapse for SAT, VAT and SKM as well as in relation to the CAC and non-CAC groups. Primary NSCLC tumours from patients recruited into TRACERx were subjected to multi-region whole-exome and RNA sequencing.

First, differential gene expression analyses were conducted according to the previously established SAT, VAT, SKM and body weight loss groups. Distinct differential gene expression profiles, adjusted for histology and sex (see methods), were observed in primary tumours from patients with $\geq 20\%$ SAT loss, $\geq 20\%$ VAT loss (Figure 3 A, B), $\geq 10\%$ SKM loss (Figure 3 C) and grade 4 BMI-adjusted weight loss (Figure 3 D), respectively. Tumours from patients with $\geq 20\%$ SAT loss showed higher expression of genes associated with cell polarity, such as *PARD3* or *MPP7* (Figure 3 A), while tumours from patients with $\geq 20\%$ VAT loss demonstrated higher expression of inflammatory genes, such as *GPR15*, or TRAIL Receptor 1 (*TNFRSF10A*) (Figure 3 B). Furthermore, tumours from patients with $\geq 10\%$ SKM loss showed higher expression of members of the Melanoma antigen Genes (MAGEA) family, including *MAGEA2* and *MAGEA3* (Figure 3 C). Tumours from patients with grade 4 BMI-adjusted weight loss were characterised by higher expression of inflammatory

genes, such as lipopolysaccharide binding protein (*LBP*), and enzymes involved in glucose metabolism, such as proprotein convertase 1 (*PCSK1*) (Figure 3 D). These distinctive expression profiles share strikingly little overlap between subgroups (Figure 3 E), suggesting the possibility of differential wasting mechanisms.

On a gene set level, tumours from patients with $\geq 20\%$ SAT/VAT loss or grade 4 weight loss shared similar increases in inflammatory signalling (TNF α , IL6-JAK-STAT3, IFN α /IFN γ and inflammatory response signalling), while tumours from patients with $\geq 10\%$ SKM loss demonstrated higher expression epithelial and tissue proliferative gene sets, such as epithelial-mesenchymal transition, hedgehog, and WNT- β -catenin signalling (Figure S5).

Next, the previously established classification into CAC and non-CAC groups was leveraged to compare differential tumour gene expression profiles between the two groups: Melanoma-Associated Antigen 6 *MAGEA6* and transcriptional and cytoskeletal regulators, such as *NR2F1* and *SPTB*, demonstrated significantly increased expression in the CAC group (Figure 3 F). In keeping with this, significantly enriched hallmark gene sets in this group included epithelial-mesenchymal transition, hedgehog signalling and myogenesis (Figure 3 I).

To explore the biological relevance of these differentially expressed genes in the development of CAC, we cross-referenced our significant outputs with a list of 400 genes (hereafter referred to as 'cachexia candidate gene list') generated by reviewing supporting literature. The genes from this list are known to be associated with the development of either obesity or cachexia in pre-clinical and clinical models, or derived from genome wide association studies of obesity, BMI, and cachexia (Table S8) {Speliotes, 2010 #1536} {Locke, 2015 #1537; Wen, 2014 #1538; Akiyama, 2017 #1539; Winkler, 2015 #1540; Johns, 2017 #1541; Solheim, 2011 #1542; Baranski,

2018 #1543;Lodge, 2021 #1544;Ding, 2021 #1545;Kwon, 2015 #1546;Song, 2019 #1547;Newton, 2020 #1548;Figuroa-Clarevega, 2015 #1549;Kim, 2021 #1550}. From this gene list, Semaphorin-3A (*SEMA3A*), Insulin-like growth factor 1 (*IGF1*), potassium channel *KCNJ12*, and Akinase Anchoring Protein 6 (*AKAP6*) were found to be differentially upregulated in the CAC group (Figure 3H).

Somatic copy number analysis was performed using GISTIC2.0 to identify amplifications and deletions specific to the SAT/VAT/SKM and body weight loss prognostic thresholds, and the CAC and non-CAC groups (Figure S6, Figure 3 I){Mermel, 2011 #1556}. Multiple loci from the cachexia candidate gene list were found to be exclusively amplified in the CAC group, including chromosome 11q22.3 containing various metalloproteinases, such as *MMP1* and *MMP3*, and chromosome 3q27.1, containing *ADIPOQ*.

Since some inflammatory signalling pathways (interferon-alpha response) were observed to be upregulated in the CAC group, we further investigated whether this was concordant with increased inflammatory cell infiltration using the previously published TCRA algorithm, providing T-cell infiltration estimates based on whole-exome sequencing data {Bentham, 2021 #1551}. No significant difference in TCRA scores, i.e., T-cell infiltrates, was observed between primary tumours in the CAC and non-CAC group, nor when adjusted for sex (Figure S7 A-C). However, TCRA scores in the blood were higher in the non-CAC group, suggesting higher circulating T-cell levels in the circulation in patients without CAC (Figure S7 D-F). In addition, we used the Danaher immune signature {Danaher, 2017 #1555} and CIBERSORTx {Newman, 2019 #1552} RNAseq cell type deconvolution approaches to compare the abundance of inflammatory cells in tumours in the CAC and non-CAC groups, but no difference was observed (Figure S8 and Figure S9).

In summary, the analysis of tumour WES and RNAseq data revealed distinct tumour genomic and transcriptomic profiles that suggest possible mechanisms and contributors to the development of CAC, including upregulated inflammatory signalling or increased expression and amplification of matrix metalloproteinases.

Associations between circulating proteins and cancer cachexia features

Given the systemic implications of CAC on the tumour host, plasma samples collected at diagnosis (141 patients: 55 in CAC, 70 non-CAC group) and at relapse (115 patients, 43 in CAC, 58 in non-CAC group) were subjected to unbiased proteomic profiling using the Olink Explore 3072 platform to investigate the presence of potential circulating mediators of CAC {Assarsson, 2014 #1514}(Figure S10). Differential plasma protein expression was observed to varying extents between the different groups of SAT, VAT, SKM and BMI-adjusted weight loss groups (Figure S11). Between the CAC and non-CAC groups, 79 proteins, including TRAIL2 receptor 2 (TNFRSF10B), EDA2R and HSPA2, were observed to be preferentially expressed in patients in the CAC group (Figure 4 A), whereas 9 proteins, including SLC28A1, MENT and PDE4D, were more abundant in patients in the non-CAC group (Table S8). Increased tumour gene expression and concordantly increased plasma protein expression was only found for *HSPA2* and *KIAA0319* in the CAC group (Table S8).

Of 24 circulating proteins previously reported to play a role in CAC{Lerner, 2016 #1451} (Table S10) only CCL11, IL5, TNF and GDF15 showed higher expression in plasma from patients with CAC. Of these, only GDF15, a known mediator of anorexia and weight loss{Lockhart, 2020 #1486}, was significantly differentially elevated following Benjamin-Hochberg correction (Figure 4B). The

normalised plasma protein expression of GDF15 was significantly higher in patients in the CAC group compared with patients in the non-CAC group ($p < 0.001$) (Figure 4 C). Furthermore, a significant correlation between normalised protein expression of GDF15 and loss of SAT, VAT, SKM and body weight particularly at relapse, was observed (Figure S12).

This relationship was confirmed using an ELISA-based GDF15 assay (Roche Elecsys GDF15) in a cohort of TRACERx patients for whom plasma at diagnosis (107 patients) and at relapse (89 patients) was available (Figure S13). The median circulating GDF15 levels were 1902 pg/ml (normal range 200-1200 pg/ml {Lockhart, 2020 #1486}) at diagnosis, and further increased at relapse (median 2393.5 pg/ml) (Figure 4 D). Notably, serum GDF15 levels in the TRACERx cohort (median age 70 years) were higher than previously published in age-matched non-cancer volunteers (60-70 years, median plasma GDF15 levels of 866 pg/ml) and in line with previous reports of increased circulating GDF15 in patients with NSCLC {Wollert, 2017 #1515; Cai, 2020 #1571; Roche, 2020 #1516; Welsh, 2022 #1601}. Median GDF15 levels of patients in the CAC group were overall significantly higher (2483 pg/ml) compared to 1756 pg/ml in the non-CAC group ($p < 0.001$) (Figure S14). Circulating GDF15 levels were significantly associated with increased age at diagnosis and the squamous cell carcinoma histology subtype. There was no significant association between GDF15 levels and baseline BMI, smoking status or number of pack years, use of adjuvant treatment, tumour stage or volume (Figure S15). GDF15 baseline and relapse levels were not associated with time to recurrence in a Cox regression analysis (HR 1.0, 95% CI 1.0-1.0, $p = 0.269$ and $p = 0.350$, respectively).

Circulating GDF15 levels, at baseline and at relapse, were significantly higher in patients with grade 4 BMI-adjusted weight loss at relapse compared to those

patients who remained stable or gained weight (grade 0) (Figure 4 E, F). Based on body composition analysis, increased circulating GDF15, particularly at relapse, was significantly associated with loss of body weight, as well as SAT, VAT and SKM tissue (Figure 4 G-J, Figure S14). Whole exome and RNA sequencing data were analysed to investigate whether increased circulating GDF15 was associated with genomic alterations and/or increased gene expression in the primary tumour. Mutations in the GDF15 gene were found in only one patient (c.A313G: p.I105V). There was no significant correlation between GDF15 gene expression and circulating GDF15 levels (Figure 4 K, L), although the power to detect any correlation was limited by the small number of relapse samples in this cohort. Furthermore, the ploidy-adjusted copy number of the GDF15 gene on chromosome 19p13.11q did not correlate with gene expression (Figure S16A) or with circulating GDF15 levels (Figure S16B). However, GDF15 copy number gains in relapsed tumour tissue, observed in 8 patients, were associated with higher circulating GDF15 levels compared to tumours with copy number losses ($p=0.045$, Figure S16C).

To explore whether the relationship between increased circulating GDF15 and the development of CAC was restricted to early-stage NSCLC, we measured plasma GDF15 levels in 164 patients with stage IV lung adenocarcinoma in an independent cohort of patients with metastatic NSCLC (NCT01360554) {Ramalingam, 2014 #1446}. Circulating GDF15 levels were significantly higher in patients with grade 3 and 4 BMI-adjusted weight loss compared to patients with stable body weight (Figure S17), suggesting a similar relationship between circulating GDF15 and weight loss at least in the advanced disease setting.

Overall, these data suggest that patients who develop features of CAC at first relapse have distinct tumour genomic and transcriptomic, as well as plasma

proteomic, profiles. Among these, circulating GDF15 showed the strongest correlation with loss of adipose tissue as well as skeletal muscle tissue and body weight, underlining the potential of GDF15-targeted therapy for reversing CAC features.

DISCUSSION

This study integrates longitudinal body composition with matched tumour whole exome and RNA sequencing data, as well as plasma proteomic data in a combined analysis of the biological correlates associated with the development of CAC. Our first aim was to generate body composition profiles of patients with early-stage lung cancer and to elucidate associations with survival outcomes. At the time of cancer diagnosis, we found that low SAT, VAT and SKM as individual measures of body composition were associated with poor LCSS and OS in the TRACERx cohort. These findings were further validated in the independent BLCS cohort and remained significant in a multivariable model for both cohorts when accounting for potential confounders, such as age and smoking status {He, 2018 #1525; Bamia, 2004 #1524}. Whilst low SAT has previously been shown to be associated with shorter OS in patients with cancer {Ebadi, 2017 #1523}, to the best of our knowledge we demonstrate here for the first time the relevance of low VAT as an independent prognostic factor, associated with both overall and lung cancer-specific survival at the time of diagnosis in early stage NSCLC.

Two distinct features of CAC are the loss of body weight and skeletal muscle {Fearon, 2011 #1444}. To incorporate this temporal dimension, and to focus on features of cachexia likely to be driven by the underlying cancer, the second aim of this study was to identify TRACERx patients who experienced disease recurrence

during follow-up after primary surgery with available CT scans and body weight for analysis. Importantly, body composition measurements were made using an established deep learning-based algorithm to avoid investigator-dependent variation in the datasets between diagnosis and relapse {Cespedes Feliciano, 2020 #1511;Dabiri, 2020 #1510}. To identify patients with and without features of CAC for further cohort-level analyses, we used specific thresholds for altered body composition in SAT ($\geq 20\%$ loss), VAT ($\geq 20\%$ loss) and SKM ($\geq 10\%$ loss) that were associated with significantly shorter LCSS and OS. Notably, most patients experienced a loss across all three body compartments, but small subgroups of patients with pronounced or isolated loss of SAT, VAT or SKM were also identified, suggesting the possibility that distinct CAC clinical phenotypes may exist whereby individual, or a combination of, body composition compartments may be preferentially affected {Sartori, 2021 #1526}.

Having stratified TRACERx patients into CAC and non-CAC groups, the third aim of the study was to explore possible molecular mechanisms and mediators of CAC in an unbiased discovery approach. To this end, we analysed the genomic and transcriptomic profiles of tumours in these respective groups to investigate the presence of alterations potentially associated with CAC, albeit restricted to the primary, as opposed to relapse, tumour. By cross-referencing the differential gene expression profiles generated from bulk tumour analysis with list of 400 genes derived from published reports relating to obesity and cachexia, we identified Semaphorin-3A (*SEMA3A*), Insulin-like growth factor 1 (*IGF1*), the potassium channel *KCNJ12*, and A-kinase anchor protein 6 (*AKAP6*) to be differentially higher expressed in tumour in the CAC group {van der Klaauw, 2019 #1531}. Additionally, we observed differential upregulation of multiple inflammatory pathways in the CAC

group and detected higher expression levels of several genes of interest, including lipopolysaccharide binding protein (*LBP*) and metallopeptidases, such as *ADAMTS3* {Bindels, 2018 #1553;Cal, 2015 #1554}. This is in line with previous observations that suggested tumour-driven systemic inflammation as a contributor to CAC {de Matos-Neto, 2015 #1582;Webster, 2020 #1635}. While several pre-clinical studies have reported increased tumour expression of a range of purported mediators of cachexia, such as IL17 in the murine Lewis Lung Carcinoma model and PTH-related protein {Ying, 2022 #1532;Kir, 2014 #1533}, we did not see significant differential expression of these genes in the primary tumours between the CAC and non-CAC groups. Furthermore, the analysis of tumour copy number profiles revealed distinct copy number alterations of patients in the CAC group. These included amplification of a cluster of MMPs on chromosome 11q22.3, including MMP1 and MMP3, which have been previously associated with tumour-induced muscle loss in drosophila models {Lodge, 2021 #1544}. Amplifications found in the CAC group also included ADIPOQ (Adiponectin) on chromosome 3q27.1. High levels of Adiponectin, a regulator of insulin sensitivity and lipid metabolism, have been previously associated with low body weight and weight loss, while low levels were associated with obesity {Achari, 2017 #1610}. Hence, these transcriptomic and genomic analyses suggest possible tumour-derived factors associated with the development of CAC.

To study the presence of potential circulating mediators of the cachexia phenotype, we profiled the plasma proteomes of 115 patients both at diagnosis and relapse. Several candidate mediators of cachexia exhibited significant differential plasma protein expression between CAC and non-CAC groups. These included MMP3, with the corresponding gene also being amplified in the tumour tissue, and the proinflammatory cytokine TNF- α . The most differentially, highly expressed,

candidate plasma protein among patients with CAC was GDF15; a highly conserved member of the Transforming Growth Factor β (TGF- β) superfamily which circulates at physiologically low levels in healthy states. GDF15 expression and secretion is upregulated in response to cellular stress {Rochette, 2020 #1461; Mulderrig, 2021 #1495}, and elevated circulating levels have been identified in a broad range of human diseases, including cardiac, renal, and respiratory failure, and notably anorexia and weight loss {Tsai, 2016 #1432; Johnen, 2007 #1442; Lerner, 2015 #1443; Tsai, 2018 #1431; Tsai, 2018 #1431}. There is mounting pre-clinical evidence that GDF15 could present a druggable target; with transgenic mice overexpressing GDF15 developing a cachexia-like syndrome that can be readily reversed with neutralising anti-GDF15 monoclonal antibodies {Johnen, 2007 #1442; Emmerson, 2017 #1434}.

We employed an orthogonal GDF15-specific serological assay, demonstrating that circulating GDF15 levels, particularly at relapse, correlated significantly with loss of SAT, VAT, SKM and body weight - suggesting a cachexia-mediating role for GDF15 in patients with NSCLC. The same association between circulating GDF15 levels and body weight loss was observed in an independent cohort of advanced metastatic NSCLC, suggesting that this association is agnostic of disease stage, and further corroborating the key role of GDF15 as a mediator of CAC in NSCLC. No clear evidence of increased tumour GDF15 expression was observed and no significant genomic amplification of the GDF15 locus was detected, however, GDF15 is known to be produced in a diversity of tissue sites, including liver and kidney tissue, which may act as potential sources of pathological secretion. {Patel, 2022 #1535; Mulderrig, 2021 #1495}. Furthermore, GDF15 is subject to differential rates of

production and clearance that are mediated by other factors, such as hepatic stabilin-1 and -2 {Schledzewski, 2011 #1572}.

While our study provides a NSCLC dataset integrating tumour genomics and plasma proteomics with body composition and body weight, there are limitations in the interpretation of the data. Firstly, the body composition analysis focuses on two timepoints in the disease course: diagnosis and first relapse. Conceivably, the cachexia phenotype may develop at subsequent time points of disease progression which warrants further investigation. Moreover, our description of the CAC phenotype is based on changes that occur between diagnosis and first relapse, whereas the tumour genomic and transcriptomic data were mostly derived from the resected primary tumour. As such, our analyses are correlative and hypothesis-generating and therefore further functional experiments are warranted. A selection bias cannot be excluded in the plasma proteomic analyses, since only patients with available complete baseline and relapse imaging data and sufficient amounts of banked plasma were analysed. Future studies aiming to establish the underlying biological mechanisms of CAC would benefit from incorporating functional data, such as physical activity, food intake and muscle function, as well as quality of life data, to reflect the complexity of CAC and to capture its related constitutional symptoms. Finally, the identified thresholds used in this study to stratify patients into CAC and non-CAC groups are applicable to our TRACERx cohort alone, and their use in other studies would require further validation.

Overall, this study demonstrates the significant, independent, and prognostic impact of altered body composition on clinical outcomes in NSCLC, including lung cancer-specific survival. The presence of specific body composition changes, either predominant loss of adipose or muscle tissue, in subgroups of patients suggests

distinct clinical subtypes of CAC, which may be driven by unique biological mechanisms warranting further investigation. We show that automated pipeline technologies unlock the potential to leverage CT imaging embedded in medical oncological practice to identify patients at risk of developing CAC, simultaneously providing the scientific means to study potential drivers and mechanisms of cachexia pathophysiology. Amidst the plethora of proposed pro-cachectic mediators, GDF15 emerges as a differentially expressed, and clinically measurable, protein, with a mounting evidence base establishing its potential to translate to a biotherapeutic target.

METHODS

Patient cohorts

TRACERx cohort

TRACERx is a UK-wide prospective multi-centre study of patients with primary NSCLC that aims to define evolutionary trajectories for lung cancer through multiregion and longitudinal tumour sampling {Jamal-Hanjani, 2017 #1318}(NCT01888601). The study was approved by an independent Research Ethics Committee (13/LO/1546). All patients provided written informed consent. Patients are followed for up to 5 years from the point of primary diagnosis, through surgical resection to cure, cancer progression(s) and death. The study collects longitudinal clinical, epidemiological, and imaging data as well as multi-region tumour tissue samples. The study protocol with inclusion/exclusion criteria has been published previously.{Jamal-Hanjani, 2017 #1318}

For the body composition and cachexia study, patients were included if a pre-operative abdominal CT incorporating the third lumbar vertebrae and performed within 3 months of primary surgical tumour resection, was available. For the delta cohort, all patients with available abdominal CT at the time of disease recurrence were included. Overall, 651 patients with a pre-surgery baseline CT were included, of which 188 had a relapse with corresponding CT scan.

BCLS cohort

Patients from the US cancer cohort in this analysis are part of the ongoing Boston Lung Cancer Study (BLCS), a multi-institutional epidemiology cohort study at MGB, and the DFCI. Inclusion criteria for this analysis were pathology confirmed diagnosis of lung cancer with available abdominopelvic CT or PET/CT scans within 4 months before and the 2 months after diagnosis, and relevant, a priori defined clinical covariates available. Patients were excluded if imaging or clinical data were incomplete or missing. The study was approved by the institutional review boards of all institutions and the requirement for written informed consent was waived. Information regarding smoking status was prospectively collected in the TRACERx study and was also available for all the patients in the BLCS cohort. Patient demographics are provided in Table 1.

ARCHER1009 cohort

ARCHER1009 was a randomised phase 3 study for patients with advanced NSCLC who were randomised to the EGFR inhibitors dacomitinib versus erlotinib.{Ramalingam, 2014 #1446} This study received ethical approval from the Pfizer Institutional Review Board and had been conducted in accordance with the

Declaration of Helsinki (ClinicalTrials.gov Identifier: NCT01360554). For the cachexia study, 164 patients with available plasma and body weight data at two timepoints within 6 months were included. Classification of weight loss was similar to published schemes {DeWys, 1981 #1447;Fearon, 2011 #1444} and defined as either weight stable/gain, >0-5% weight loss, and >5% weight loss. Any prior treatment (chemotherapy, radiation, or surgery) must have been completed at least 2 weeks prior to randomization at the start of the study.

Body composition and body weight measurements

For the TRACERx cohort, subcutaneous adipose tissue, visceral adipose tissue, and skeletal muscle tissue were quantified as areas (cm²) localised to the third lumbar vertebrae (L3) level. L3 selection, image segmentation and tissue area quantification were conducted via an automated deep learning-based pipeline (DAFS platform, Voronoi Health Analytics, Vancouver, British Columbia, Canada {Popuri, 2016 #1602;Ma, 2021 #1603;Cespedes Feliciano, 2020 #1511;Dabiri, 2020 #1510}). To measure SAT, VAT and SKM areas at the L3 level, analyses were run with the 'avg-L3mid[3]' command, which measures across 3 slices above and below the midst of L3 in order to increase data accuracy (Figure S18). The following Hounsfield unit (HU) boundaries were used: For SAT -190 to -30 HU, for VAT -150 to -50 HU and for SKM -29 to 150 HU. Accurate L3 selection and segmentation quality was manually inspected for all patients by an experienced investigator (medical oncologist with over 10 years experience). To this end, each CT annotation and segmentation was inspected via a sagittal, coronal and axial view of each scan using the 'quickcheck' option; mis-annotations was corrected using the CAST (CT Annotation and Segmentation Tool) feature from DAFS. To confirm the highly

reproducible nature of the algorithm, 60 CT-scans were re-run twice through the automated annotation, segmentation and measurements steps. Using the precision metrics previously proposed by Arribas et al, $\leq 0.01\%$ variance was observed between the runs (Table S12) {Arribas, 2022 #1611}. Absolute (cm^2) change between baseline and relapse was calculated for SAT, VAT and SKM per patient.

For the independent BLCS cohort, baseline body composition was measured by a previously developed automated system {Bridge, 2018 #1512;Magudia, 2021 #1513}, which produced body composition measurements at the L3 level for SAT, VAT and SKM in cm^2 . Manual quality control was conducted for all images in the BLCS cohort by an experienced investigator (radiologist with over 10 years experience) .

Body weight data in the TRACERx, BCLS and ARCHER1009 cohorts were collected from medical records in routine health care settings; scale calibrations and type of clothing was conducted according to local standards. For the TRACERx cohort, body weight changes were assessed between baseline and relapse; for the ARCHER1009 cohort, body weight changes were assessed between time of enrolment and end of study. BMI-adjusted weight change was calculated as follows, similar to previous publications {Martin, 2014 #1454}:

	Body mass index (kg/m^2)				
Weight loss (%)	≥ 28	25–27.9	22–24.9	20–21.9	<20
± 2.4	0	0	1	1	3
2.5–5.9	1	2	2	2	3
6–10.9	2	3	3	3	4
11–14.9	3	3	3	4	4
≥ 15	3	4	4	4	4

Whole exome and RNA sequencing

Whole exome sequencing was performed on DNA purified from tumour tissue with matched germline from whole blood, as described previously {Jamal-Hanjani, 2017 #1318; Rosenthal, 2019 #1324}. RNA was extracted from primary tumour tissue, and downstream analyses were conducted as reported previously; where applicable, multi-region gene expression data were handled as average per tumour or adjusted by a linear-mixed effects model, also accounting for histology and sex (Martinez-Rui et al, 2023). P value <0.01 and absolute log fold change >1 was used to identify differentially expressed genes. Gene set enrichment analysis (GSEA) was done by pre-ranking the genes by fold-change and using the fgsea R package (v1.22.0) and using the hallmark gene set from the Molecular Signatures Database (v.7.4) with a minimal size of 15 and maximal size of 500. GISTIC2.0 was used to analyse copy number alterations in tumour tissues according to cachexia and non-cachexia group {Mermel, 2011 #1556}.

A list of candidate cachexia genes, i.e. genes possibly associated with cachexia, was generated by reviewing supporting literature, including genome wide association studies on obesity and cachexia, since both extremes can share common metabolic perturbations, as well as murine and drosophila models (see Table S8).

Plasma proteomics

Plasma proteomes were profiled using the Olink Explore 3072 platform (Olink, Uppsala, Sweden) at Bioxpedia (Aarhus, Denmark) following the standard Olink-certified protocol.{Assarsson, 2014 #1514} TRACERx plasma samples and control samples were plated on three 96 well plates that were processed in one

batch. For data analysis, protein expression as log₂ of normalized protein expression (NPX) was used. Comparisons of protein expression between body composition groups were made by Welch 2-sample t-tests with Benjamin-Hochberg correction. Data that did not pass the Olink-specified quality control metrics were excluded from the analysis. R packages OlinkAnalyze v3.1.0 and ggplot2 v3.3.6 were used for data analysis and visualization.

For the orthogonal GDF15 validation, GDF15 was measured in TRACERx plasma samples using the Elecsys GDF15 immunoassay (F. Hoffmann-La Roche, Basel, Switzerland) at baseline ahead of surgery and at diagnosis of relapse. Plasma samples from the ARCHER1009 samples were obtained at the end of study and human GDF15 was measured using ELISA (R&D Systems, Minneapolis, MN).

Statistical analyses

All statistical analyses were performed in R v4.0.3. GDF15 plasma levels were log₁₀ transformed. Tests involving correlations were done using `stat_cor` from the R package `ggpubr` (v04.0) with linear regression and Spearman's rank-order correlation. Categorical comparisons were made by ANOVA, student's t-test, or Wilcoxon test. P-values were two-sided and considered as statistically significant if below 0.05. Plotting and analysis in R was also done by `ggplot2` (v3.3.3), `cowplot` (v1.1.1), `tidyr` (v1.1.1) and `dplyr` (v1.0.4).

TABLES

Table 1. Baseline patient characteristics of TRACERx and BLCS cohorts.

SD standard deviation, BMI body mass index, VAT visceral adipose tissue, SAT subcutaneous adipose tissue, SKM skeletal muscle

	TRACERx (N=651)	BLCS (N=420)
Age , Mean (SD), years	68.7 (9.12)	64.53 (10.1)
BMI , mean (SD), kg/m ²	26.7 (5.11)	26.48 (5.4)
Weight , mean (SD), m	75.0 (16.9)	75.02 (18.2)
Height , mean (SD), m	1.67 (0.1)	1.68 (0.1)
VAT baseline , mean (SD), cm ²	138 (98.7)	145.5 (113.9)
SAT baseline , mean (SD), cm ²	169 (94.3)	190.5 (103.5)
SKM baseline , mean (SD), cm ²	126 (34.5)	133.6 (33.9)
Sex		
Female	286 (43.9%)	198 (47.1)
Male	365 (56.1%)	222 (52.9)
Ethnicity		
White-Caucasian		411 (97.9)
African American		9 (2.1)
Non-White	36 (5.5%)	
White-British-Irish	590 (90.6%)	
White-Other	22 (3.4%)	
Missing	3 (0.5%)	
Smoking status		
Current Smoker	89 (13.7%)	155 (36.9)
Ex-Smoker	513 (78.8%)	216 (51.4)
Never Smoked	49 (7.5%)	49 (11.7)
NSCLC stage		
IA	131 (20.1%)	203 (48.3) - all stage I
IB	139 (21.4%)	
IIA	124 (19.0%)	63 (15.0) - all stage II
IIB	113 (17.4%)	
IIIA	138 (21.2%)	154 (36.7) - all stage III
IIIB	6 (0.9%)	
Histology		

	TRACERx (N=651)	BLCS (N=420)
Adenocarcinoma	362 (55.6%)	223 (53.1)
Squamous cell carcinoma	211 (32.4%)	104 (24.8)
Other	65 (10.0%)	93 (22.1)
N/A	13 (2.0%)	
Adjuvant treatment		
Adjuvant	227 (34.9%)	229 (54.5)
No adjuvant	408 (62.7%)	191 (45.5)
Missing	16 (2.5%)	

FIGURES

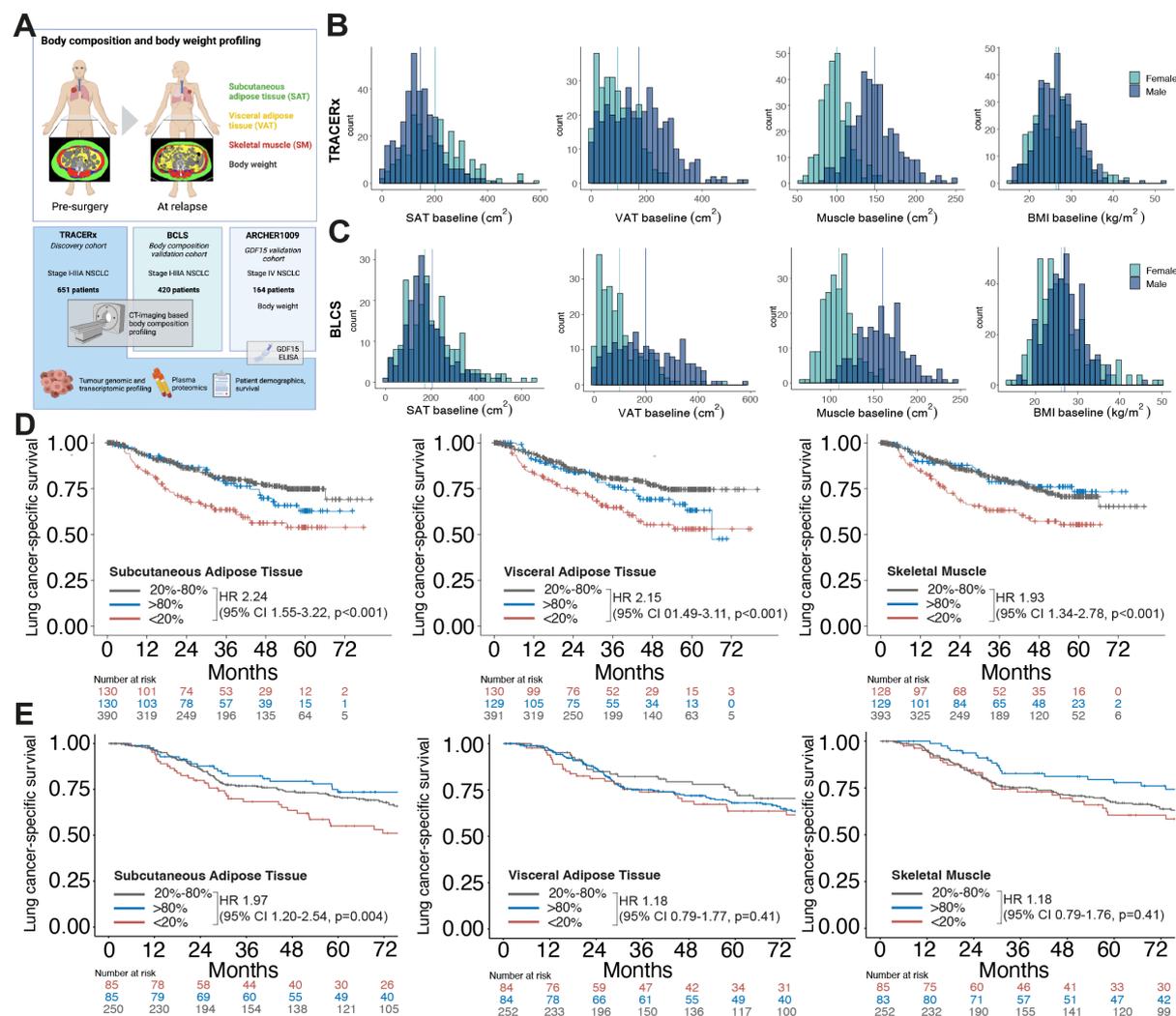


Figure 1. Body composition and cancer-specific survival in the TRACERx and BLCS studies.

A Study outline of body composition and downstream analyses. **B** Distribution of body composition and body mass metrics according to sex at primary diagnosis in

the TRACERx cohort (N=651) and **C** BLCS cohort (N=420); vertical lines indicate mean in female (green) and male (blue) patients. **D** Lung cancer-specific survival in TRACERx and **E** BLCS cohorts for body composition at primary diagnosis according to sex-specific bottom 20th percentile (red, <20%), mid 60th percentile (grey, 20%-80%) and upper 20th percentiles (blue, >80%) of subcutaneous adipose tissue (SAT), visceral adipose tissue (VAT) and skeletal muscle (SKM). Hazard ratios and 95% confidence intervals (CI) derived from univariate Cox regression analysis of the <20% group and the 20-80% groups.

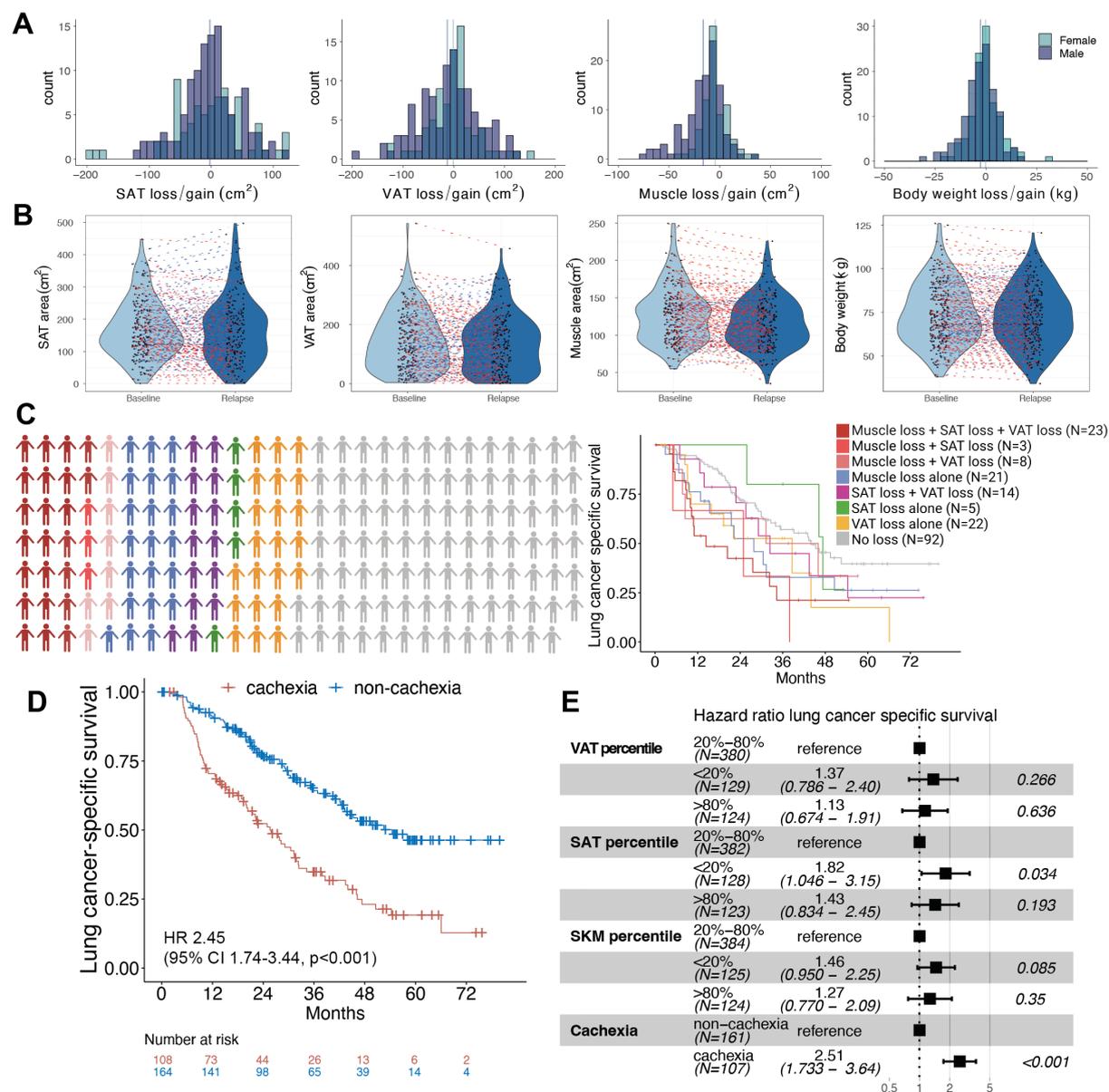


Figure 2. Survival outcomes according to changes in body composition between primary diagnosis and first relapse.

A Distribution of losses and gains of body composition and body weight in the TRACERx cohort between primary diagnosis and first tumour recurrence according to sex (N=188). B Dynamics of body composition (N=188) and body weight (N=232) between primary diagnosis (“Baseline”) and first recurrence (“Relapse”), with red

lines indicating decrease/loss, blue lines indicating increase/gain, and grey line indicating no changes. **C** Subgroups of patients according to (co-)presence of SAT loss (defined as $\geq 20\%$ between baseline and relapse), VAT loss ($\geq 20\%$) and SKM loss ($\geq 10\%$). Kaplan-Meier analysis of lung cancer-specific survival according to labelled subgroups. **D** Lung cancer-specific survival according to cachexia and non-cachexia status. **E** Multivariable analysis of lung cancer-specific survival, adjusted for age, sex, BMI, smoking status, disease stage, histological subtype, ethnicity, and adjuvant therapy.

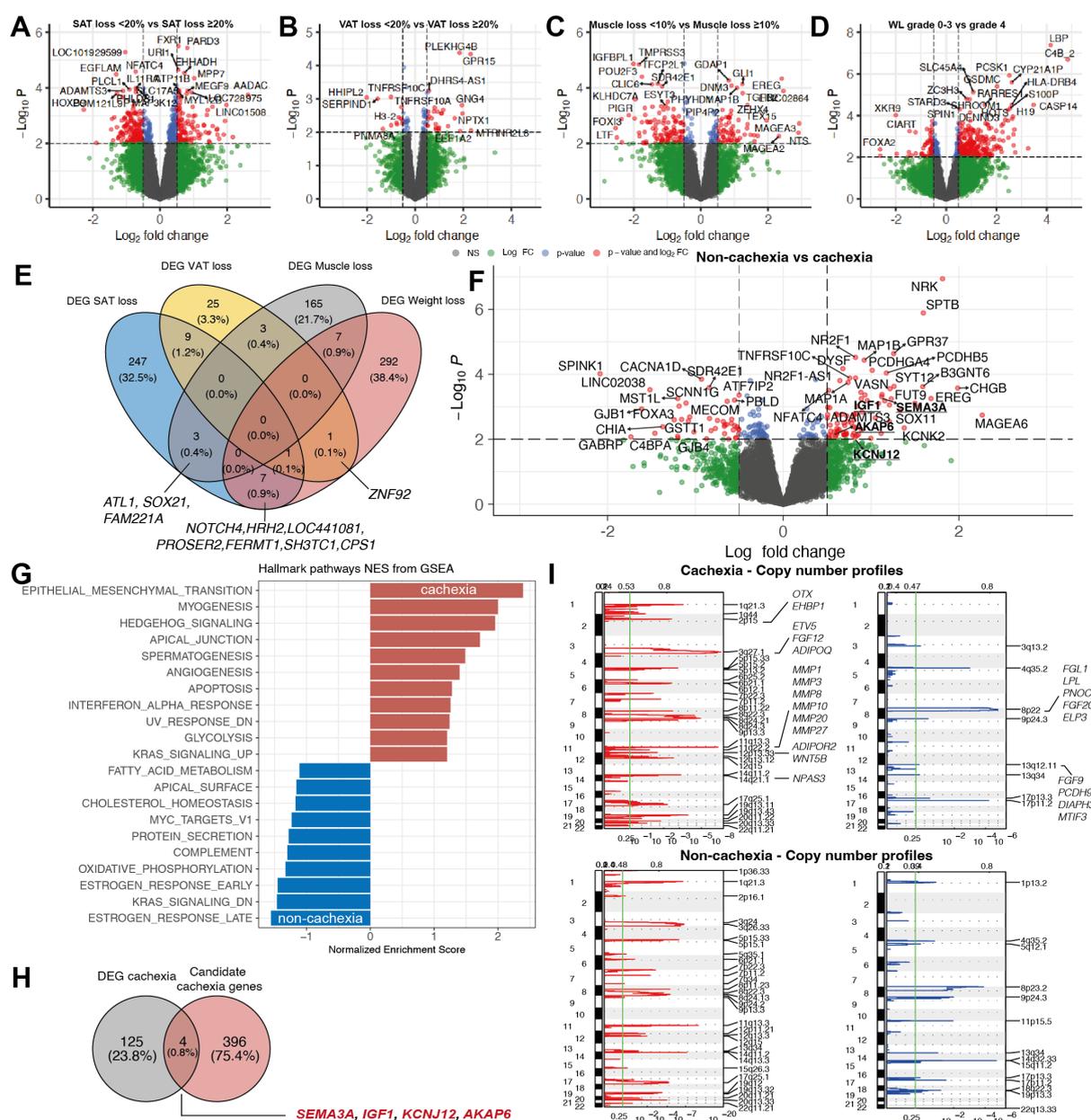


Figure 3. Tumour genomic and transcriptomic profiles according to body composition and cancer cachexia.

Tumour differential gene expression between patients with **A** SAT loss <20% versus $\geq 20\%$, **B** VAT loss <20% versus $\geq 20\%$, **C** Muscle loss <10% versus $\geq 10\%$, **D** BMI adjusted weight loss grade 0-3 versus 4, all adjusted for number of tumour regions, sex, and histology. **E** Overlap of differentially expressed genes (DEG) between the

$\geq 20\%$ SAT, $\geq 20\%$ VAT, $\geq 10\%$ SKM and grade 4 weight loss groups. **F Tumour differential gene expression between patients in the non-cachexia versus cachexia group** adjusted for number of tumour regions, sex, and histology. **G** Hallmark gene set enrichment in the cachexia (red) versus non-cachexia groups (blue), adjusted for sex and histology. **H Overlap of differentially expressed genes between the cachexia group and candidate cachexia genes.** **I** GISTIC analysis of copy number alterations of cachexia (upper row) and non-cachexia groups (lower row). Y-axes indicate chromosomal positions (1-22), red plots indicate gains, blue plots indicate losses. X-axes indicate q-values. Most significant peaks are indicated on the right of each panel; regions with FDR $q \leq 0.25$ (vertical green line) are considered significant. GISTIC G-Scores are plotted on top of each panel.

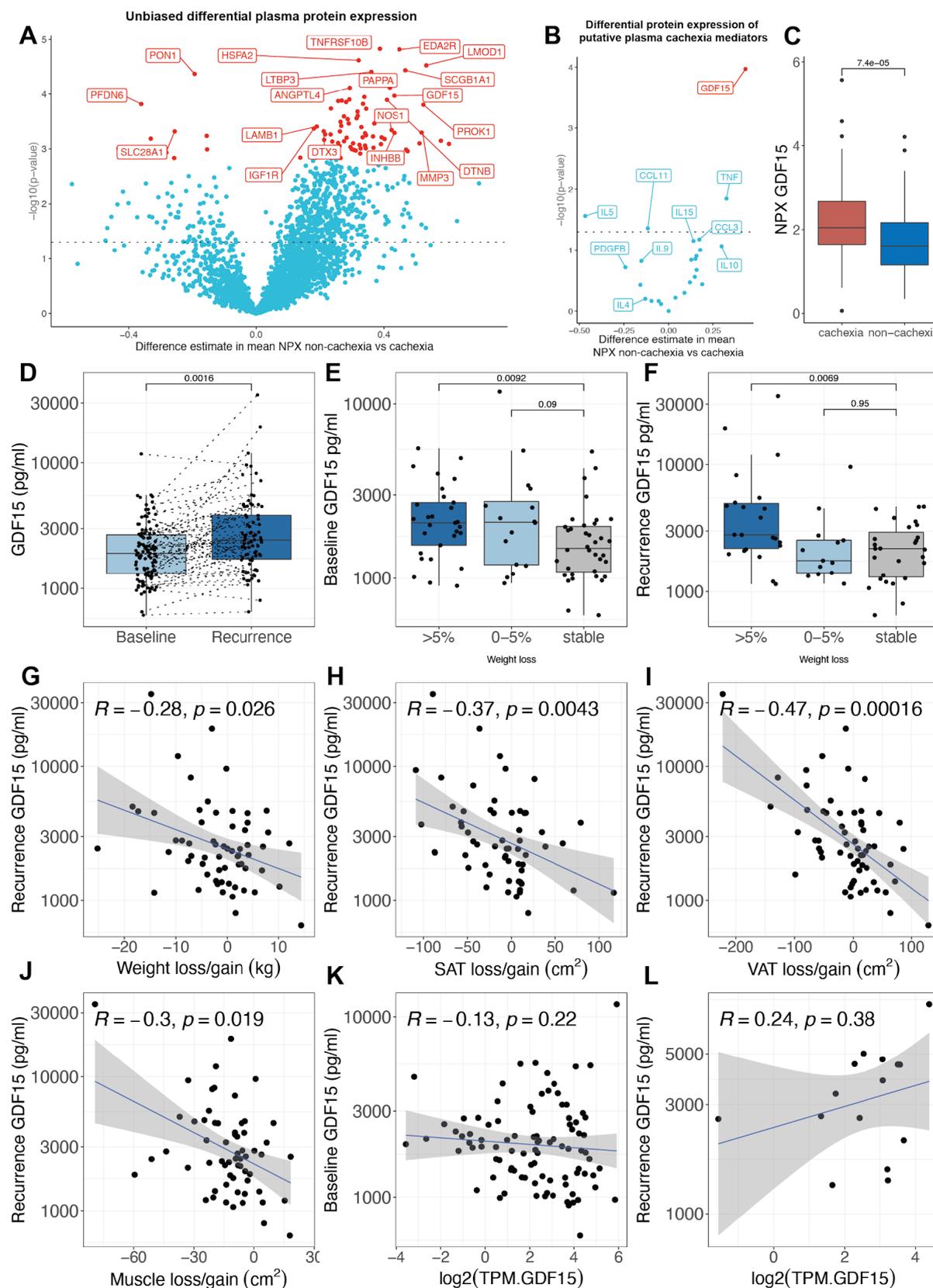


Figure 4. Differential protein expression and associations between circulating GDF15, body composition and body weight changes, and cancer cachexia. A Differential plasma proteome of patients in the non-cachexia versus cachexia groups. Red labels indicate significant differential protein expression after adjusting

for multiplicity. **B** Differential plasma protein expression of putative cachexia mediators in the non-cachexia versus cachexia group. **C** Normalized plasma protein expression of GDF15 in the non-cachexia versus cachexia groups (two-sided Wilcoxon test). **D** Plasma GDF15 levels in patients at diagnosis (baseline) or first recurrence of NSCLC in the TRACERx cohort (two-sided Wilcoxon test). **E** Baseline and **F** recurrence GDF15 levels according to weight change category in the TRACERx cohort (two-sided Wilcoxon test, error bars indicate standard deviation). **G-J** Spearman correlation of recurrence GDF15 levels and loss/gain of body weight (N=62)(**G**), subcutaneous adipose tissue (SAT)(**H**), visceral adipose tissue (VAT) (**I**), and muscle (**J**)(N=61). **K-L** Spearman correlation of baseline (**K**) and recurrence (**L**) GDF15 levels and GDF15 gene expression as transcripts per million (TPM). All Wilcoxon tests are two-sided and box plots represent lower quartile, median and upper quartile, whiskers extend to a maximum of $1.5 \times$ IQR beyond the box. Points indicate individual data points. Grey shade areas represent 95% confidence intervals. Y-axes represent log₁₀ scales.

ACKNOWLEDGMENTS

We gratefully acknowledge the patients and their families, the investigators and site personnel who participated in the TRACERx and ARCHER1009 study. TRACERx is funded by Cancer Research UK (C11496/A17786) and coordinated through the Cancer Research UK and UCL Cancer Trials Centre. ARCHER1009 was funded by Pfizer.

Figure 1A was created with BioRender.com.

OAS is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 467697427.

MJH has received funding from Cancer Research UK, National Institute for Health Research, Rosetrees Trust, UKI NETs and NIHR University College London Hospitals Biomedical Research Centre.

TBKW is supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001169), the UK Medical Research Council (FC001169) and the Wellcome Trust (FC001169) as well as the Marie Curie ITN Project PLOIDYNET (FP7-PEOPLE-2013, 607722), Breast Cancer Research Foundation (BCRF), Royal Society Research Professorships Enhancement Award (RP/EA/180007) and the Foulkes Foundation.

CM-R is supported by the Rosetrees Trust (M630).

NM is a Sir Henry Dale Fellow, jointly funded by the Wellcome Trust and the Royal Society (Grant Number 211179/Z/18/Z) and receives funding from Cancer Research UK, Rosetrees and the NIHR BRC at University College London Hospitals and the CRUK University College London Experimental Cancer Medicine Centre.

SOR is supported by the Medical Research Council MRC.MC.UU.12012.1, a Wellcome Senior Investigator Award 214274/Z/18/Z and the NIHR Cambridge Biomedical Research Centre.

CS is Royal Society Napier Research Professor (RP150154). CS is funded by Cancer Research UK (TRACERx, PEACE and CRUK Cancer Immunotherapy Catalyst Network), Cancer Research UK Lung Cancer Centre of Excellence (C11496/A30025), the Rosetrees Trust, Butterfield and Stonegate Trusts, NovoNordisk Foundation (ID16584), Royal Society Professorship Enhancement Award (RP/EA/180007), the National Institute for Health Research (NIHR) Biomedical Research Centre at University College London Hospitals, the Cancer Research UK-University College London Centre, Experimental Cancer Medicine Centre, and the Breast Cancer Research Foundation (BCRF 20-157).

DATA AND CODE AVAILABILITY STATEMENT

The RNA sequencing and Whole Exome Sequencing data (from the TRACERx study) generated, used or analysed during this study are available through the Cancer Research UK & University College London Cancer Trials Centre (ctc.tracex@ucl.ac.uk) for academic non-commercial research purposes only, subject to review of a project proposal that will be evaluated by a TRACERx data

access committee and any applicable ethical approvals, and entering into an appropriate data access agreement. All code to reproduce figures will be made available at publication or upon request from reviewers.

AUTHOR DISCLOSURE INFORMATION

OAS: Advisory Board (AstraZeneca, AbbVie, Ascentage, Gilead, Janssen, Roche), speaker honoraria (Adaptive, AstraZeneca, AbbVie, BeiGene, Eli Lilly, Gilead, Janssen, Roche), research funding (BeiGene, AbbVie, Janssen, Roche)

JML: Receives research funding from the National Institute for Health and Care Research (NIHR).

SC: Employment (Pfizer)

MR: Receives research funding from the National Institutes of Health, Lustgarten Foundation, Stand Up to Cancer Foundation, and the Hale Family Center for Pancreatic Cancer at Dana-Farber Cancer Institute.

NS: Consultancy for Amgen, Astrazeneca, Boehringer Ingelheim, Eli-Lilly, Hanmi Pharmaceuticals, Novartis, Novo Nordisk, Pfizer, and Sanofi. Research grants for Astrazeneca, Boehringer Ingelheim, Roche Diagnostics, and Novartis

PW: Grant income from Roche Diagnostics, Astrazeneca, Boehringer Ingelheim, and Novartis and speaker's fees from Novo Nordisk outside the submitted work.

KP: KP actively directs Voronoi Health Analytics Incorporated, a Canadian corporation that sells commercial licences for the DAFS (Data Analysis Facilitation Suite) software.

MFB: MFB actively directs Voronoi Health Analytics Incorporated, a Canadian corporation that sells commercial licences for the DAFS (Data Analysis Facilitation Suite) software.

NM: Stock options, consultancy (Achilles Therapeutics)

DMB: Employment (Pfizer)

SOR: Consultancy (Pfizer, AstraZeneca, Northsea, Third Rock Ventures)

NJB: is a co-inventor to a patent to identify responders to cancer treatment (PCT/GB2018/051912), and a patent to predict HRD deficiency (US10190160B2)

MJH: M.J-H is a CRUK Career Establishment Awardee and has received funding from CRUK, IASLC International Lung Cancer Foundation, Lung Cancer Research Foundation, Rosetrees Trust, UKI NETs, NIHR, NIHR UCLH Biomedical Research Centre. M.J-H. has consulted and is a member of the Scientific Advisory Board and Steering Committee, for Achilles Therapeutics, has received speaker honoraria from AstexPharmaceuticals, Oslo Cancer Cluster, and holds a patent PCT/US2017/028013 relating to methods for lung cancer detection.

CS: Research grants (Pfizer, AstraZeneca, Bristol Myers Squibb, Roche-Vertana, Boehringer-Ingelheim, Archer Dx Inc, Ono Pharmaceutical, MSD); Consultancy (Pfizer, AstraZeneca, Bristol Myers Squibb, Roche-Vertana, MSD, GRAIL, Medicxi, Bicycle Therapeutics, Metabomed, Sarah Cannon Research Institute, Illumina, Genentech, Apogen Biotechnologies, Epic Bioscience, Achilles Therapeutics, Novartis, Amgen, GlaxoSmithKline).

All other authors have nothing to disclose.

REFERENCES

1. Martin, L., *et al.* Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **31**, 1539-1547 (2013).
2. Calle, E.E., Rodriguez, C., Walker-Thurmond, K. & Thun, M.J. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *The New England journal of medicine* **348**, 1625-1638 (2003).
3. Guo, Y., *et al.* Body mass index and mortality in chronic obstructive pulmonary disease: A dose-response meta-analysis. *Medicine (Baltimore)* **95**, e4225 (2016).
4. Lavie, C.J., McAuley, P.A., Church, T.S., Milani, R.V. & Blair, S.N. Obesity and cardiovascular diseases: implications regarding fitness, fatness, and severity in the obesity paradox. *J Am Coll Cardiol* **63**, 1345-1354 (2014).

5. Baracos, V.E., Martin, L., Korc, M., Guttridge, D.C. & Fearon, K.C.H. Cancer-associated cachexia. *Nature Reviews Disease Primers* **4**, 17105 (2018).
6. Fearon, K., *et al.* Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol* **12**, 489-495 (2011).
7. Shen, W., *et al.* Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol (1985)* **97**, 2333-2338 (2004).
8. Shen, W., *et al.* Visceral adipose tissue: relations between single-slice areas and total volume. *Am J Clin Nutr* **80**, 271-278 (2004).
9. Mourtzakis, M., *et al.* A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl Physiol Nutr Metab* **33**, 997-1006 (2008).
10. Caan, B.J., *et al.* Association of Muscle and Adiposity Measured by Computed Tomography With Survival in Patients With Nonmetastatic Breast Cancer. *JAMA Oncol* **4**, 798-804 (2018).
11. Lee, J.S., *et al.* Subcutaneous Fat Distribution is a Prognostic Biomarker for Men with Castration Resistant Prostate Cancer. *J Urol* **200**, 114-120 (2018).
12. Baracos, V.E., Reiman, T., Mourtzakis, M., Gioulbasanis, I. & Antoun, S. Body composition in patients with non-small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am J Clin Nutr* **91**, 1133s-1137s (2010).
13. Yang, M., Shen, Y., Tan, L. & Li, W. Prognostic Value of Sarcopenia in Lung Cancer: A Systematic Review and Meta-analysis. *Chest* **156**, 101-111 (2019).
14. Popinat, G., *et al.* Sub-cutaneous Fat Mass measured on multislice computed tomography of pretreatment PET/CT is a prognostic factor of stage IV non-small cell lung cancer treated by nivolumab. *Oncoimmunology* **8**, e1580128 (2019).
15. Tan, H., *et al.* Preoperative Body Composition Combined with Tumor Metabolism Analysis by PET/CT Is Associated with Disease-Free Survival in Patients with NSCLC. *Contrast Media Mol Imaging* **2022**, 7429319 (2022).
16. Jamal-Hanjani, M., *et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *The New England journal of medicine* **376**, 2109-2121 (2017).
17. Christiani, D. Boston Lung Cancer Study. (<https://sites.sph.harvard.edu/blcs/>, 2022).
18. Pavai, D.R., *et al.* A systematic review examining the relationship between cytokines and cachexia in incurable cancer. *J Cachexia Sarcopenia Muscle* **13**, 824-838 (2022).
19. Loumaye, A., *et al.* Role of Activin A and Myostatin in Human Cancer Cachexia. *The Journal of Clinical Endocrinology & Metabolism* **100**, 2030-2038 (2015).
20. Hsu, J.-Y., *et al.* Non-homeostatic body weight regulation through a brainstem-restricted receptor for GDF15. *Nature* **550**, 255-259 (2017).
21. Dabiri, S., *et al.* Deep learning method for localization and segmentation of abdominal CT. *Comput Med Imaging Graph* **85**, 101776 (2020).
22. Bridge, C.P., *et al.* Fully-Automated Analysis of Body Composition from CT in Cancer Patients Using Convolutional Neural Networks. in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis* (eds. Stoyanov, D., *et al.*) 204-213 (Springer International Publishing, Cham, 2018).
23. Martin, L., *et al.* Diagnostic Criteria for the Classification of Cancer-Associated Weight Loss. *Journal of Clinical Oncology* **33**, 90-99 (2014).
24. Speliotes, E.K., *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature genetics* **42**, 937-948 (2010).
25. Locke, A.E., *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
26. Wen, W., *et al.* Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Mol Genet* **23**, 5492-5504 (2014).
27. Akiyama, M., *et al.* Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nature genetics* **49**, 1458-1467 (2017).

28. Winkler, T.W., *et al.* The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet* **11**, e1005378 (2015).
29. Johns, N., *et al.* New genetic signatures associated with cancer cachexia as defined by low skeletal muscle index and weight loss. *J Cachexia Sarcopenia Muscle* **8**, 122-130 (2017).
30. Solheim, T.S., *et al.* Is there a genetic cause for cancer cachexia? - a clinical validation study in 1797 patients. *Br J Cancer* **105**, 1244-1251 (2011).
31. Baranski, T.J., *et al.* A high throughput, functional screen of human Body Mass Index GWAS loci using tissue-specific RNAi *Drosophila melanogaster* crosses. *PLOS Genetics* **14**, e1007222 (2018).
32. Lodge, W., *et al.* Tumor-derived MMPs regulate cachexia in a *Drosophila* cancer model. *Dev Cell* **56**, 2664-2680.e2666 (2021).
33. Ding, G., *et al.* Coordination of tumor growth and host wasting by tumor-derived Upd3. *Cell Rep* **36**, 109553 (2021).
34. Kwon, Y., *et al.* Systemic organ wasting induced by localized expression of the secreted insulin/IGF antagonist Impl2. *Dev Cell* **33**, 36-46 (2015).
35. Song, W., *et al.* Tumor-Derived Ligands Trigger Tumor Growth and Host Wasting via Differential MEK Activation. *Dev Cell* **48**, 277-286.e276 (2019).
36. Newton, H., *et al.* Systemic muscle wasting and coordinated tumour response drive tumourigenesis. *Nature Communications* **11**, 4653 (2020).
37. Figueroa-Claresvega, A. & Bilder, D. Malignant *Drosophila* tumors interrupt insulin signaling to induce cachexia-like wasting. *Dev Cell* **33**, 47-55 (2015).
38. Kim, J., *et al.* Tumor-induced disruption of the blood-brain barrier promotes host death. *Dev Cell* **56**, 2712-2721.e2714 (2021).
39. Mermel, C.H., *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biology* **12**, R41 (2011).
40. Bentham, R., *et al.* Using DNA sequencing data to quantify T cell fraction and therapy response. *Nature* **597**, 555-560 (2021).
41. Danaher, P., *et al.* Gene expression markers of Tumor Infiltrating Leukocytes. *Journal for ImmunoTherapy of Cancer* **5**, 18 (2017).
42. Newman, A.M., *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773-782 (2019).
43. Assarsson, E., *et al.* Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability. *PLOS ONE* **9**, e95192 (2014).
44. Lerner, L., *et al.* MAP3K11/GDF15 axis is a critical driver of cancer cachexia. *J Cachexia Sarcopenia Muscle* **7**, 467-482 (2016).
45. Lockhart, S.M., Saudek, V. & O'Rahilly, S. GDF15: A Hormone Conveying Somatic Distress to the Brain. *Endocr Rev* **41**(2020).
46. Wollert, K.C., *et al.* An Automated Assay for Growth Differentiation Factor 15. *The Journal of Applied Laboratory Medicine* **1**, 510-521 (2017).
47. Cai, D., *et al.* Extensive serum biomarker analysis in patients with non-small-cell lung carcinoma. *Cytokine* **126**, 154868 (2020).
48. Roche, H.-L. Elecsys GDF-15 Datasheet. (2020).
49. Welsh, P., *et al.* Reference ranges for GDF-15, and risk factors associated with GDF-15, in a large general population cohort. *Clin Chem Lab Med* (2022).
50. Ramalingam, S.S., *et al.* Dacomitinib versus erlotinib in patients with advanced-stage, previously treated non-small-cell lung cancer (ARCHER 1009): a randomised, double-blind, phase 3 trial. *The Lancet Oncology* **15**, 1369-1378 (2014).
51. He, X., *et al.* Age- and sex-related differences in body composition in healthy subjects aged 18 to 82 years. *Medicine* **97**(2018).
52. Bamia, C., Trichopoulou, A., Lenas, D. & Trichopoulos, D. Tobacco smoking in relation to body fat mass and distribution in a general population sample. *International Journal of Obesity* **28**, 1091-1096 (2004).
53. Ebadi, M., *et al.* Subcutaneous adiposity is an independent predictor of mortality in cancer patients. *Br J Cancer* **117**, 148-155 (2017).

54. Cespedes Feliciano, E.M., *et al.* Evaluation of automated computed tomography segmentation to assess body composition and mortality associations in cancer patients. *J Cachexia Sarcopenia Muscle* **11**, 1258-1269 (2020).
55. Sartori, R., Romanello, V. & Sandri, M. Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. *Nature Communications* **12**, 330 (2021).
56. van der Klaauw, A.A., *et al.* Human Semaphorin 3 Variants Link Melanocortin Circuit Development and Energy Balance. *Cell* **176**, 729-742.e718 (2019).
57. Bindels, L.B., *et al.* Increased gut permeability in cancer cachexia: mechanisms and clinical relevance. *Oncotarget* **9**, 18224-18238 (2018).
58. Cal, S. & López-Otín, C. ADAMTS proteases and cancer. *Matrix Biology* **44-46**, 77-85 (2015).
59. Ying, L., *et al.* IL-17A contributes to skeletal muscle atrophy in lung cancer-induced cachexia via JAK2/STAT3 pathway. *American Journal of Physiology-Cell Physiology* **322**, C814-C824 (2022).
60. Kir, S., *et al.* Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia. *Nature* **513**, 100-104 (2014).
61. Rochette, L., Zeller, M., Cottin, Y. & Vergely, C. Insights Into Mechanisms of GDF15 and Receptor GFRAL: Therapeutic Targets. *Trends in Endocrinology & Metabolism* **31**, 939-951 (2020).
62. Mulderrig, L., *et al.* Aldehyde-driven transcriptional stress triggers an anorexic DNA damage response. *Nature* **600**, 158-163 (2021).
63. Tsai, V.W., Lin, S., Brown, D.A., Salis, A. & Breit, S.N. Anorexia-cachexia and obesity treatment may be two sides of the same coin: role of the TGF- β superfamily cytokine MIC-1/GDF15. *Int J Obes (Lond)* **40**, 193-197 (2016).
64. Johnen, H., *et al.* Tumor-induced anorexia and weight loss are mediated by the TGF- β superfamily cytokine MIC-1. *Nature Medicine* **13**, 1333-1340 (2007).
65. Lerner, L., *et al.* Plasma growth differentiation factor 15 is associated with weight loss and mortality in cancer patients. *J Cachexia Sarcopenia Muscle* **6**, 317-324 (2015).
66. Tsai, V.W., *et al.* Treatment with the TGF- β superfamily cytokine MIC-1/GDF15 reduces the adiposity and corrects the metabolic dysfunction of mice with diet-induced obesity. *Int J Obes (Lond)* **42**, 561-571 (2018).
67. Emmerson, P.J., *et al.* The metabolic effects of GDF15 are mediated by the orphan receptor GFRAL. *Nature Medicine* **23**, 1215-1219 (2017).
68. Suriben, R., *et al.* Antibody-mediated inhibition of GDF15–GFRAL activity reverses cancer cachexia in mice. *Nature Medicine* **26**, 1264-1270 (2020).
69. Bernardo, B., *et al.* Characterization of cachexia in the human fibrosarcoma HT-1080 mouse tumour model. *J Cachexia Sarcopenia Muscle* **11**, 1813-1829 (2020).
70. Patel, S., *et al.* Endogenous GDF15 and FGF21 additively alleviate hepatic steatosis and insulin resistance in obese mice. *bioRxiv*, 2022.2006.2008.495255 (2022).
71. Schledzewski, K., *et al.* Deficiency of liver sinusoidal scavenger receptors stabilin-1 and -2 in mice causes glomerulofibrotic nephropathy via impaired hepatic clearance of noxious blood factors. *J Clin Invest* **121**, 703-714 (2011).
72. Breen, D.M., *et al.* GDF-15 Neutralization Alleviates Platinum-Based Chemotherapy-Induced Emesis, Anorexia, and Weight Loss in Mice and Nonhuman Primates. *Cell Metabolism* **32**, 938-950.e936 (2020).
73. Mullican, S.E., *et al.* GFRAL is the receptor for GDF15 and the ligand promotes weight loss in mice and nonhuman primates. *Nature Medicine* **23**, 1150-1157 (2017).
74. Yang, L., *et al.* GFRAL is the receptor for GDF15 and is required for the anti-obesity effects of the ligand. *Nature Medicine* **23**, 1158-1166 (2017).
75. DeWys, W.D., Costa, G. & Henkin, R. Clinical parameters related to anorexia. *Cancer Treatment Reports* **65**, 49-52 (1981).
76. Popuri, K., Cobzas, D., Esfandiari, N., Baracos, V. & Jägersand, M. Body Composition Assessment in Axial CT Images Using FEM-Based Automatic Segmentation of Skeletal Muscle. *IEEE Trans Med Imaging* **35**, 512-520 (2016).
77. Ma, D., Chow, V., Popuri, K. & Beg, M.F. Comprehensive Validation of Automated Whole Body Skeletal Muscle, Adipose Tissue, and Bone Segmentation from 3D CT images for

- Body Composition Analysis: Towards Extended Body Composition. *ArXiv abs/2106.00652*(2021).
78. Magudia, K., *et al.* Population-Scale CT-based Body Composition Analysis of a Large Outpatient Population Using Deep Learning to Derive Age-, Sex-, and Race-specific Reference Curves. *Radiology* **298**, 319-329 (2021).
 79. Rosenthal, R., *et al.* Neoantigen-directed immune escape in lung cancer evolution. *Nature* **567**, 479-485 (2019).