Title

Immune system homeostasis in people with multiple long-term conditions determines susceptibility to organ injury and mortality following cardiac surgery

Authors: *Florence Y Lai MPhil,1 *Adewale S Adebayo PhD,1 *Sophia Sheikh PhD,1 Marius Roman MD,1 Lathishia Joel-David BSc,1 Hardeep Aujla MSc,1 Tom Chad MBChB,1 Kristina Tomkova PhD,1 Shameem Ladak PhD,1 Gianluigi Condorelli MD,2 Mustafa Zakkar PhD,1 Charles Solomon PhD1, Marcin J Woźniak PhD,1 Gavin J Murphy MD.1

*Contributed equally.

Author Affiliations:

1Department of Cardiovascular Sciences and NIHR Cardiovascular Biomedical Research Centre, University of Leicester, Glenfield Hospital, Leicester, LE3 9QP, UK.
2Department of Cardiovascular Medicine, Humanitas Research Hospital, Humanitas University, Milan.

Corresponding Author: Professor Gavin Murphy, Department of Cardiovascular Sciences, University of Leicester, Clinical Sciences Wing, Glenfield General Hospital, Leicester, LE3 9QP, UK. Email: gjm19@leicester.ac.uk

Word Count: 4997

Keywords: Cardiovascular surgery, organ protection, inflammageing
ABSTRACT

**Background:** People with Multiple Long-Term Conditions (MLTC) experience higher rates of organ failure and death following cardiac surgery. The aim of this study was to identify disease processes in MLTC associated with increased susceptibility to organ injury post-surgery.

**Methods:** Unsupervised machine learning methods were applied to pre-surgery biomarkers of haematological, cardiac, liver, and renal disease from four intensively phenotyped cardiac surgery cohorts. The resulting MLTC clusters were then matched to a fifth cohort where single nuclei RNA sequencing (snRNAseq) was performed on atrial biopsies collected at surgery.

**Results:** K-means clustering identified two MLTC clusters. Cluster 1 had lower rates of chronic kidney disease and anaemia and increased immune system activation pre-surgery. Cluster 2 had more severe cardiorenal disease, anaemia, and elevated biomarkers of immunological ageing pre-surgery. Cluster 2 had significantly higher rates of organ injury relative to Cluster 1. The results were consistent across internal and external validation analyses.

Analysis of snRNAseq data in biopsies from Cluster 1 demonstrated enrichment for immune response genes in cardiomyocytes, naive T/B lymphocytes and progenitor cells, and activation of non-tissue resident macrophages relative to Cluster 2. Cluster 2 showed enrichment for senescent/effector memory T cells, dysregulated activation of tissue-resident macrophages, and cardiomyocyte dedifferentiation relative to Cluster 1.

In UK Biobank, genetic modification of genes differentially expressed between the two MLTC phenotypes altered 90-day mortality post-surgery.

**Conclusions:** Immune system homeostasis determines susceptibility to organ injury and death in people with MLTC undergoing cardiac surgery and represents a previously unrecognised target for organ protection interventions.
Introduction

After cardiac surgery, people with Multiple Long-Term Conditions (MLTC), also referred to as multimorbidity, (1) demonstrate increased susceptibility to organ injury affecting the heart, lungs, brain, or kidneys, higher mortality, and slower functional recovery.(2) Clinical trials evaluating perioperative interventions targeting ischaemia reperfusion, the systemic inflammatory response, endothelial homeostasis, or blood transfusion have failed to identify effective organ protection interventions.(3-6) However, over 95% of variation in outcomes post-cardiac surgery between operators or institutions is attributable to age and long-term conditions at baseline.(7, 8) It follows that disease processes associated with MLTC may offer novel targets for pre-surgery organ protection interventions. The mechanisms underlying MLTC are not well characterised. The aim of this study was to use individual patient data and myocardial single nuclei transcriptomics data from six existing cardiac surgery cohorts, and informatics approaches to identify disease processes in MLTC phenotypes at baseline that alter susceptibility to organ injury following cardiac surgery.

Methods

Study design

A study workflow diagram is provided in Figure 1. UK ethics approvals for the primary studies were 09/H0104/53, 12/EM/0475, 13/EM/0383, 15/YH/0489, and 16/NW/0494, and, and for the secondary clustering analyses and UK Biobank analyses were 18/WS/0148, 21/NW/0157 and 21/SC/0118. The clustering analyses were pre-specified (see eSupplement and (9)) albeit with knowledge of results from previous analyses of the datasets.(3, 10, 11) The results are reported as per the STrengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.(12)

Data Sources

The development cohort included three studies that used cohort enrichment tools for MLTC, (4, 13) and performed detailed phenotyping of baseline conditions, post-surgery organ injury, and platelet/leucocyte activation in blood: 1. A case-control study to identify the role of microvesicle and microvesicle-derived micro-RNA in post-cardiac surgery acute kidney injury (MaRACAS).(14) 2. A RCT of red cell washing for the prevention of transfusion-associated organ injury in cardiac surgery (REDWASH) (15) and 3. A RCT to study the effect of intravenous sildenafil citrate on post-cardiac surgery acute kidney (REVAKI-2).(16) These studies were undertaken across 3 UK centres, used common outcome definitions, had low levels of missing data, and performed biomarker and flow cytometry analyses in a single laboratory using standardised methods.(8) External validation was performed in participants of the Coagulation and platelet laboratory testing in cardiac surgery (COPTIC) study (11) with a
Leicester AKI score >22,(4) as used in the MARACAS and REVAKI-2 studies. The single
nuclei RNA sequencing (snRNAseq) analyses used data from the Case control study to
identify the role of epigenetic regulation of genes responsible for energy metabolism and
mitochondrial function in the Obesity Paradox in cardiac surgery (ObCARD) study.(10)
Transcriptomic data was externally validated in the Imputed Genotype Data v3 from UK
Biobank (Application 77596).

**Input variables for cluster analysis**

MLTC was defined as the presence of two or more long-term health conditions.(17)

Input variables in the development cohort included continuously distributed baseline patient
characteristics and pre-surgery serum biomarkers of long-term conditions common to all three
studies. These were: Age and body mass index (BMI), serum bilirubin and Glomerular filtration
rate (eGFR) estimated from serum creatinine, platelet count, and haemoglobin concentration
(Hb), transferrin, ferritin, serum urokinase-type Plasminogen Activator Receptor (SuPAR)
(suPARnostic®, ViroGates, Birkerød, Denmark), interleukins (IL)-6 (EASIA™ kit, Invitrogen),
IL-8 (Bender Medsystem GmbH, Vienna, Austria), serum Neutrophil Gelatinase Associated
Lipocalcin (NGAL), high sensitivity serum Troponin I (Enzo® Troponin I (human) ELISA kit,
Enzo Life Sciences, Ann Arbor, MI, USA) and pro-NT BNP (Human ProcartaPlex™ Simplex
assay, Thermo Fisher Scientific, Waltham, MA, USA), all measured on the DS2® 2-Plate
ELISA Processing System (Dynex Technologies, Chantilly, VA). All input variables were
numeric.

Based on the relative contributions of input variables to clustering in the development cohort,
input variables in the COPTIC cohort were limited to age, body mass index, platelet count,
Bilirubin, Haemoglobin, eGFR, IL-6, IL-8, Troponin I, Ferritin, and Transferrin.

**Outcomes**

Clinical outcomes were restricted to those measured prospectively in the individual studies.

**In the development cohort** post-operative acute kidney injury (AKI) was defined using
KDIGO criteria (18) and urinary tissue inhibitor of metalloproteinases-2 (TIMP-2)*insulin-like
growth factor-binding protein 7 (IGFBP7) (all EIA kit, EKF diagnostics, Cardiff, UK), acute lung
injury (ALI) determined using Berlin criteria, (19) myocardial injury defined as peak serum
hsTroponin I >5000ng.ml⁻¹ within 24 hours postoperatively, and systemic inflammatory
response syndrome as defined clinically,(20) Assessments of platelet and leucocyte activation
at baseline common to all three studies in the development cohort included flow cytometry
(Cyan ADP; Beckman Coulter) in whole blood for Platelet Activating Complex (PAC)-1 (BD
CD41 (Affymetrix, Santa Clara, CA, USA) and leucocyte activation (CD64, CD163; Affymetrix).
Microvesicle concentrations were calibrated using the NanoSight NS500 (Malvern
Pananalytical), with surface markers/ cell derivation identified using flow cytometry as described previously.(21) Baseline platelet function was assessed using the Multiplate® analyzer (Roche, Rotkreuz, Switzerland).

In the COPTIC study, clinical outcomes were in-hospital death, KDIGO defined acute kidney injury, myocardial infarction, stroke, sepsis, transfusion, and bleeding, all as defined in the primary publication.(11) Analyses for validated plasma protein biomarkers of biological and immunological ageing at baseline were performed using Luminex® Multiplex Assays (Luminex, Genk, Belgium) on the Magpix® platform (Luminex). These included CCL11/Eotaxin, CX3CL1/Fractalkine, GDF-15, IL5, CXCL9/MIG, CCL3/MIP-1 alpha, high sensitivity C Reactive Protein, and Serpin. PhenoAge, a validated biomarker of biological ageing, was calculated from routinely collected laboratory data.(22) Two components of the score; red cell distribution width and serum glucose data were not available and were imputed to generate a modified PhenoAge for analysis. Coagulation was phenotyped as reported in the primary report of the COPTIC study.(11)

In the ObCARD study snRNAseq was performed as described using biopsies (30 – 100 mg) collected prior to cardiopulmonary bypass from the right atrial appendage, (23) immediately snap-frozen in liquid nitrogen, and stored at -80°C. After nuclei isolation, cDNA amplification and library construction were performed using Chromium Single Cell 3’ Gel-bead in Emulsion (GEM) and Chromium Single Cell 3’ Library kits v3.1. (#1000128, 10x Genomics, Pleasanton, United States) following the manufacturer’s instructions.(24) The resulting cDNA was quality-checked using the Agilent Bioanalyzer High Sensitivity DNA kit (Agilent Technologies, Santa Clara, United States). Libraries were constructed, including fragmentation, end-repair, A-tailing, and adaptor ligation steps. A unique sample index was added to each sample. Following library construction, samples were quality-checked, and libraries were purified and eluted using Buffer EB (#19086, Qiagen, Hilden, Germany), and SPRIselect beads (#B23317, Beckman Coulter, Pasadena, United States). Samples were sequenced paired-end on Illumina NovoSeq at 2.5x108 reads per sample and analysed using existing pipelines as described below. Reads in fastq formats were QC’ed and processed using 10X Genomics CellRanger v6.0.1. Internal validation of snRNAseq data using RNAscope and immunofluorescence (Bio-Techne Ltd.) were performed as described previously.{McQueen, 2023 #805}

Statistical analysis

We applied k-means clustering to the development and validation cohorts independently using the R-package cluster. Input variables were log-transformed if skewed. Pearson’s correlation coefficient was calculated to identify any highly correlated variables. Variables that are highly correlated (Pearson’s correlation >0.6) were filtered. Missing data in input variables were low (range 0% - 8.5%). To maximise the inclusion of patient data, missing data were imputed.
using random forest imputation (R package \textit{missForest}). Variables were scaled to the standard normal distribution with a mean of 0 and a standard deviation of 1 prior to clustering. The optimal number of clusters was determined based on Silhouette distance. We additionally considered the NbClust program in R (25) and used the majority rule of the 23 indices of cluster cohesiveness and separation to confirm the optimal cluster numbers.

For internal validation in the development cohort, we conducted agglomerative hierarchical clustering analysis in the development cohort using the Ward's minimum variance method. After the clusters were identified, we characterised the clinical phenotypes based on participants' baseline characteristics and their post-op organ injury outcomes. Differences in characteristics across phenotypes were compared using the Wilcoxon test for continuous variables and the chi-square test for categorical variables.

For the identification of clusters in the ObCARD data we used a logistic regression model with stepwise selection to develop a classification model to determine cluster membership. Variables were restricted to those available in this study; age, BMI, platelet, bilirubin, haemoglobin, eGFR, troponin, ferritin, transferrin, IL6 and IL8.

**Single nuclei data and patient cluster assignment** was available for 20 participants from the OBCARD study.(10)

RNA expression data were imported and further processed with Seuratv4.0 (26) following the methods of Ma et al.(27) In brief, nuclei with more than 5% of reads being mitochondrial-derived genes and less than 200 genes were removed, the count data was SCTransformed using SCTransform in Seurat, and doublets removed using doubletFinder (28) with the doubletFinder\_v3 function in a homotypic model. Multiple samples were then integrated with PrepSCTIntegration, FindIntegrationAnchors and IntegrateData functions in Seurat which utilises the most variable genes as integration anchors, and cell clustering of the integrated assay was done with UMAP algorithm (29) FindClusters function in Seurat with resolution set at 0.5 using the first 20 principal components after visualization of cumulative standard deviation of 30 principal components with ElbowPlot function. Subclusters within cell types were determined using FindSubCluster function with the integrated snn graph and 0.2 resolution.

Cell type identification was carried out by combining the 683 canonical gene markers from Ma et al (27) and automated identification using SingleR (30) with heart data from Litinukova et al (31) as the reference. Analysis was run without a cluster showing disagreement in both methods. Cell clusters were removed if detected in only a few individuals.

To further identify subtypes of major cell types such as cardiomyocytes and immune cell subtypes, gene signatures of cell states and subtypes were obtained from Molecular Signatures (MSig) DB, (32) Martini et al, (33) Litinukova et al, (31) Tucker et al, (34) and
Koenig et al.(35) For cardiomyocytes and immune cells where these markers did not clarify the subtypes, genes with at least 1.5 times higher detection in a subcluster versus all other subclusters and found in more than 40% of the subcluster were used as cluster signature.

**Differential expression and Pathway analysis:** For cell types of interest derived from cell clustering, differential genes between patient clusters were determined with FindMarkers in Seurat with default settings of minimum 10% detection in either group, log fc threshold 0.25 and adjusted p value 0.05 using normalized RNA assay in Seurat. Pathway analysis were carried out using Metascape (36) with GO Biological Process annotation sets only and q value set at 0.05. Where cells assigned to a patient cluster group far outweigh cells assigned to another group, random down sampling of the bigger group was compared to the result of the complete case analysis. This reduced heterogeneity for the bigger group but helped to confirm overall reliability of results.

**Causal Inference analysis:** UK Biobank was used to infer whether the observed differences in gene expression between clusters could influence clinical outcomes. SNPs identified as expression quantitative trait loci (eQTLs) were used as instruments for causal inference analysis. Effect sizes of SNP on expression (summary statistics) were obtained from MRbase (37, 38) which were derived from analysis on healthy tissue. Imputed genotype data v3 was obtained from UK Biobank. Per chromosome bgen files were converted and combined to genotype matrix using qctool (39), VariantAnnotation (40) and snpStats (41) packages. Major cardiac surgery procedures were defined using OPCS4 codes (K40–K46, K49, K50, K75). Outcomes were defined as death within 90 days of surgery or survival post-surgery using the death register. Effect sizes on outcomes were calculated after adjusting for age, sex and 10 principal components. Annotations of UKBiobank SNPs were obtained from NealeLab data (http://www.nealelab.is/uk-biobank) round 2 results. Both summarized data on exposure (gene expression) and outcomes were combined in an MR analysis using the TwoSampleMR package.(38) Where only single instruments were available for the exposure the Wald ratio test was applied. Where multiple instruments were available for exposure, the inverse variance weighted method was applied. The functional annotation of significant genes was obtained from GeneCards (42) and PantherDB.(43, 44)

**Code and Data Availability** The authors declare that the data supporting the findings of this study are available within the paper and its Supplementary Information. All analysis were carried out in Unix, Python and R. A list of packages and modules used in each of these environments is attached as Appendix I/ Supp Table. Scripts are attached as an Appendix. Raw single nuclei data are available upon request from the corresponding author.
Results

Development Cohort

Unsupervised K-means clustering analysis in the development cohort (n=277, Table 1) identified 2 clusters with distinct baseline characteristics (Figure 2A, B and sTables 1 and 2). Cluster 1 had a higher proportion of males (87% versus 55%), lower platelet counts and serum levels of pro-NT BNP, Troponin I, SuPAR, and IL-6, and higher bilirubin, Haemoglobin, Ferritin, and Transferrin Saturations, all measured at baseline, versus Cluster 2.

Mean age, BMI, prevalence of diabetes mellitus or previous myocardial infarction at baseline, type of operation, duration of cardiopulmonary bypass, and proportion of participants with MLTC were not statistically significantly different between the Clusters.

Cluster 1 had a lower prevalence of anaemia at baseline (37% versus 85%) versus Cluster 2. Cluster 1 also demonstrated higher levels of activated CD64+/CD163+ leucocytes, CD14+/CD41+ monocyte-platelet aggregates, and innate (CD284) and adaptive (CD3) immune cell-derived microvesicles in blood, higher levels of the canonical endothelial activation marker ICAM-1, and increased platelet reactivity at baseline. In contrast, Cluster 2 had significantly higher serum IL-6 levels and levels of, lower levels of leucocyte and endothelial activation and platelet reactivity in blood versus Cluster 1 (sTable 3).

The two clusters demonstrated different frequencies of organ injury post-surgery (Figure 2C). Cluster 2 had a higher prevalence of Stage 2/3 AKI (19% versus 6%), myocardial injury (68% versus 52%), acute lung injury (34% versus 23%), and Multiple (>2) Organ Dysfunction (MODS, 82% versus 59%) versus Cluster 1. Cluster 2 also demonstrated higher peak Troponin I and IL-6 concentrations (sTable 4).

Internal validation using an agglomerative hierarchical clustering (HC) algorithm with 2 clusters showed similar characteristics to the phenotypes identified by K-means clustering (sTables 5 and 6, and sFigure 1).

External validation

We repeated K-means clustering using the most predictive biomarkers from the development analysis in 800 participants from the COPTIC cohort. (sTable 7).

The results in the COPTIC cohort were consistent with the results in the development cohort (Figure 2D and sTable 7). Cluster 1 had a higher proportion of males (88%), haemoglobin, bilirubin, ferritin, and eGFR, at baseline versus Cluster 1. Cluster 2 had higher concentrations of serum troponin, platelet counts, transferrin concentrations, and serum IL-8 at baseline relative to Cluster 1. In contrast to the development cohort, IL-6 levels were not significantly different between the groups.
At baseline, Cluster 1 had lower rates of anaemia, CKD (46% versus 74%), NYHA Class II/IV symptoms (45% versus 52%) and cerebrovascular disease (9% versus 14%) versus Cluster 2. In contrast to the clusters in the development cohort, in COPTIC, Cluster 2 was significantly younger and had lower BMI measures/ levels of obesity versus Cluster 1.

Cluster 2 had significantly higher plasma biomarker levels of immunological ageing at baseline (Table 7), including CCL11/ Eotaxin, CX3CL1/ Fractalkine, GDF-15, and CXCL9/ MIG (iAge) and modified PhenoAge, a biomarker of biological ageing.

The COPTIC clusters demonstrated differences in rates of organ injury (Figure 2D and Table 8). Cluster 2 demonstrated higher rates of AKI (70% versus 57%) but not severe AKI (22% versus 21%), stroke (2.9% versus 0.2%), clinically important bleeding (37% versus 30%), and large volume (≥5 units) blood transfusions (15% versus 5%) post-surgery. Cluster 2 also demonstrated a dysregulated clotting phenotype post-surgery (Table 8).

Myocardial single nuclei transcriptomic analysis of the ObCARD cohort.

Next, based on the best predictors from the development cohort analyses, and restricted to biomarkers measured in the ObCARD cohort, we developed a prediction model to identify the two clusters in 20 participants from the ObCARD study where snRNAseq data was available from right atrial biopsies. The prediction model showed excellent prediction accuracy of 93% (Table 9) and classified 13 ObCARD participants in Cluster 1 and 7 in Cluster 2 (Table 10).

Patient Clusters demonstrate different cardiomyocyte and cardiac immune cell subtypes

Automated cell clustering from integrated sctransformed snRNAseq data (n=22.3 x 10^3 nuclei) produced seventeen clusters, that grouped into 11 major cell types including cardiomyocytes, fibroblasts, endothelial cells, adipocytes and cardiac immune cells (Figure 3A, B). To explore the differences between patient Clusters in immune cell activation and myocardial injury observed post-surgery in the Development and COPTIC cohorts, the snRNAseq analyses focused on cardiomyocytes (n=8719 nuclei) and cardiac immune cells (n=2065 nuclei). There was some sample heterogeneity with a median 464 (range 170-1754) nuclei per sample.

The major cell types showed distinct patterns of transcription and high correlation for canonical markers (Figure 2, Table 11). Major cell types were subclustered using predefined criteria as described above:

Cardiomyocytes subclustered into 4 subtypes (labelled CM_0-4, resolution=0.2) (Figure 3C, Table 11). CMs_0 cardiomyocytes were identified by increased expression of ALCAM, MKRN2OS, FRMD4A and BCL2. CMs_0 had a higher relative proportion of cells predicted to be in S and G1 phases, and greater expression of immune response genes compared to others (Figure 3A, B). CMs_1 cardiomyocyte expressed a de-differentiated phenotype with higher levels of GALNT18 and ENO4 which have functions in protein modification,
glycosylation and glycolysis. CMs_2 cardiomyocytes were not effectively distinguished with the set criteria but had lower levels of C8orf37-AS1, ZNF331, ZNF564 and SLC45A4, which have functions in regulation of gene expression or sucrose transport, compared to all others. CMs_3 cardiomyocytes had higher levels of TCAP, ACTC1, DES, and TPM1, consistent with a more differentiated contractile phenotype (sTable 11, sFigure 3).

The CMs_0 subtype was dominant in Patient Cluster 1, whereas the de-differentiated CMs_1 subtype was dominant in Patient Cluster 2. The contractile CM_3 subtype was not different between the groups (Figure 3D).

Cardiac lymphoid cells expressed higher levels of T Cell markers than B Cell markers and are referred to henceforth as T Cells. They subclustered into 3 subtypes (labelled TCells_0-2, resolution=0.2) (sFigure 4A, sTable 11). TCells_0 had higher levels of RMDN3, SH2D3C, MLLT1, CCBE1, SEC11C, C15orf40. TCells_1 had higher levels of SLC22A15, ABTB2, DNAJC6, SPHKAP, and SUGCT, that re-enriched in progenitor cells and naïve T cell phenotypes. TCells_2 mainly expressed CD4, CD8A and CD8B and had the highest signature of memory and effector_memory (senescent) T Cells; CLMN, RBM4, DPYD and SRGN (sFigure 4B, C).

Patient Cluster 1 demonstrated higher proportions of TCells_0 and TCells_1. Patient Cluster 2 demonstrated a higher proportion of effector_memory TCells_2 (Figure 3E).

Cardiac myeloid cells subclustered into three macrophage sub-types (labelled Macrophage_0-2, sTable 11). Macrophage_0 cells had higher levels of AKTIP, NEB, ELP5, MIR762HG and COA5 while Macrophage_1 cells had higher levels of CXCL12 and COL16A1. Macrophages_0 and _1 both expressed CCR2. Macrophage_2 cells had extremely low detection of CCR2 and mainly expressed gene markers of tissue-resident macrophages; KBTBD12, HS3ST1, MAMDC2-AS, EMP1, and were enriched for IL-6 regulated (activated) genes (sFigure 5A-D, sTable 11).

Gene signature and identification analyses did not match to conventional macrophage M1/M2 phenotypes (sFigure 5E). Macrophage_0 cells had higher expression of anti-inflammatory pathways versus Macrophage_1 & _2 cells. Conversely, expression of genes involved in cytokine production did not differ between Macrophage_0 and Macrophage_1 but was higher in Macrophage_2. Macrophage_2 and to a lesser extent macrophage_1 expressed more IL4R, CX3CR1 and CSF1R (sFigure 5F, G).

Patient-Cluster 1 demonstrated higher relative proportions of CCR2+ macrophage_0 cells. In contrast Patient Cluster 2 demonstrated higher proportions of activated CCR2- tissue resident macrophage_2 cells (Figure 3E).

Patient Clusters demonstrate different cardiomyocyte and immune cell transcripts
The numbers of nuclei in each Patient Cluster were different with 5952 and 2955 CM nuclei, and 1762 and 262 immune cell nuclei in Patient Clusters 1 and 2 respectively (sFigure 6A). Random down sampling was performed for the primary differential expression/ pathway analyses to generate comparable numbers of nuclei in Patient Cluster 1 and Patient Cluster 2. After down sampling there were 2955 CM nuclei and 287 immune cell nuclei in each Patient Cluster. The relative proportions in each cell subtype were preserved (sFigure 6B).

For cardiomyocytes, 1394 genes showed differential expression between Patient Cluster 1 and Patient Cluster 2 (Figure 4A, sTable 12). Upregulated genes in Cluster 1 versus Cluster 2 were enriched in RNA processing including cytoplasmic translation & splicing, IL-12 signaling through JAK-STAT, leukocyte activation & differentiation, cell apoptosis, cell adhesion, regulation of immune response and cell cycle processes, while downregulated genes were enriched in cardiac muscle development, differentiation, conduction & contraction, and nerve axon development (sTable 13). Top among significantly down regulated genes for cardiomyocytes in Cluster 1 versus Cluster 2 included MALAT1, a translation regulator, and MYL7 which codes for myosin light chain (sFigure 7A). Analyses of cell subtypes suggested that the differences between Clusters were mainly derived from differences in CM_0 cardiomyocyte gene expression with 95% overlap for down regulated genes and 70% overlap for upregulated genes in CM_0 versus all cardiomyocytes (sFigure 7B-D, sTable 13).

For cardiac immune cells, differential analysis was limited to the major cell types on the basis of relatively small numbers of nuclei in each of the 6 characterised subtypes.

For Lymphoid/ T Cells, 69 genes showed differential expression between Patient Cluster 1 and Patient Cluster 2 (Figure 4B, sTable 12). Upregulated genes in Cluster 1 versus Cluster 2 mainly enriched in ribosomal translation and peptide biosynthesis while downregulated genes were enriched in cardiac muscle morphogenesis, differentiation & conduction; angiogenesis; GTPase signal transduction, the electron transport chain, MAPK/ERK cascade, Wnt signaling, ion transport, programmed cell death, transcription and response to growth factors & stress (Figure 5B, sTable 13).

For myeloid cells, 3022 genes showed differential expression between Patient-Cluster 1 and Patient Cluster 2 (Figure 4C, sTable 12). Upregulated genes in Cluster 1 versus Cluster 2 mainly enriched in peptide and mRNA metabolism, translation, leukocyte activation/ differentiation, immune response activation, response to stress, OXPHOS, NADH complex assembly, apoptotic signalling, macromolecule processing & transport, NF-kB signalling and cell cycle process while downregulated genes were mainly enriched in heart muscle morphogenesis & contraction, angiogenesis, signal transduction, cell motility/projection/migration, response to growth factor/hormone/oxidative stress/hypoxia, axonogenesis, Wnt signalling, MAPK /ERK/SMAD signalling, cellular transport and proliferation (Figure 4C, sTable 13).
Sensitivity Analyses

In sensitivity analyses the GO enrichment was repeated in all nuclei. Data on differentially expressed genes and enriched pathways for all nuclei in Cluster 1 and Cluster 2 are presented in sTable 14 and sTable 15. The overlap in enriched GO terms with the primary analysis was 85-95% for all the major cell types.

Cells contributing to the results varied among individuals (sFigure 8A) and from the main results, randomising 100 cardiomyocytes and 50 immune cells per sample reduced the number of differentially expressed genes in cardiomyocytes (sFigure 8B) with an overlap of 66% for the top significant genes, and 59% in total. The overlap was 63% and 81% for myeloid and lymphoid cells respectively (sTable 14). In this sub-analysis, cardiac development, contraction & conduction, cardiac differentiation, cell adhesion/migration, axonogenesis and GO terms were replicated in CMs as well as peptide modification, mRNA metabolism, leukocyte activation, response to cytokine/hormone stimulus and immune response GO terms in macrophages (sTable 16). A further sensitivity analysis excluding two ObCARD participants with <2 multiple long-term conditions yielded similar results (sFigure8C).

Cardiac immune cells expressed genes commonly identified with cardiomyocytes and pathways involved in cardiac muscle morphogenesis. To confirm the accuracy of our annotation we cross-referenced major immune cell types identified in this analysis with published canonical markers (www.proteinatlas.org/humanproteome/tissue+cell+type) and retested as per sFigure 9 and sTable 17. All these analyses supported our initial classification of immune cells.

Validation of snRNAseq data

The results of RNAscope and combined RNAscope immunofluorescence staining were consistent with the snRNAseq data (Figure 5A).

Regulatory network analysis compared potential controllers of the differentially expressed genes between Patient Cluster 1 and Patient Cluster 2 for CMs. This identified 39 transcription factors including FOXN2/J3/P2, SOX5/6, MEF2C/2A, TBX5/20, GATA4/6 (sTable 18, Figure 5B) and their networks, including transcription, response to hormones (including glucocorticoids)/ hypoxia/ nutrients, cell differentiation and proliferation, cardiac/ myeloid/ endothelial cell differentiation, miRNA transcription, and receptor signalling (sTable 19). There were more active regulons/networks in Patient Cluster 2 (Figure 5C). The factors at the hub of these regulons had functions in immunoregulation, cell survival, and aging (sTable 19).

To infer whether interventions to modulate genes differentially expressed between Patient Cluster 1 and Patient Cluster 2 affected clinical outcomes, we used causal inference analysis
in 11,026 participants of UK Biobank who had undergone cardiac surgery of whom 187 had died within 90 days of surgery.(13) This showed that the genetically predicted expression of 9 genes from the cardiomyocyte snRNAseq analyses that demonstrated significantly different expression between clusters with fold change>1, including GMDS-DT, HLA-B, PPP1R9A, RPL3, RPS2, RYR2, S100A8, S100A9, and SPTBN1, were significant determinants of 90-day mortality (p<0.05) (Figure 5D, and sTable 20). These 9 genes had functions in the immune response, ribosomal translation, cellular calcium ion homeostasis, cytoskeletal organisation and cell cycle progression/differentiation.

Discussion

Main findings

Unsupervised clustering analysis of people with MLTC undergoing cardiac surgery using baseline biomarkers of chronic disease identified two distinct MLTC phenotypes with different patterns of immune cell activation and differential susceptibility to organ injury.

Patient Cluster 1 had less severe cardiorenal disease, lower rates of anaemia, increased canonical endothelial activation and leucocyte and platelet activation in blood, versus Cluster 2. In snRNAseq analyses, Patient Cluster 1 had a greater prevalence of cardiomyocytes expressing immune response genes, and larger proportions of naive T lymphocytes, lymphoid progenitor cells, and CCR2+ macrophages versus Cluster 2. Differentially regulated genes in Patient Cluster 1 versus Cluster 2 favoured cytoplasmic RNA processing, leukocyte activation and differentiation, cell apoptosis, cell adhesion, regulation of immune responses and cell cycle processes. These changes are consistent with homeostatic inflammation in chronic disease.(45)

Patient Cluster 2 had more advanced cardiorenal disease and anaemia, higher IL-6, and lower immune cell, endothelial cell, and platelet activation at baseline. Cluster 2 developed more severe acute kidney, lung, and myocardial injury post-surgery versus Cluster 1 in the development cohort. This most likely reflected increased susceptibility to injury given that key determinants of operative stress; operation complexity and cardiopulmonary bypass duration, were not statistically different between clusters. Cluster 2 also demonstrated attenuated leucocyte activation in the development cohort, and elevated biomarkers of immunological ageing in the COPTIC cohort. Cluster 2 over-representation of effector memory (senescent) T lymphocytes and dysregulated tissue resident macrophage (CCR2-) activation in RNAseq analyses are consistent with Inflammageing/ Immunosenescence phenotypes described in experimental models of immune ageing.(46)

Clinical Importance
We suggest that the results are consistent with the immune system exhaustion hypothesis of inflammageing, (45) where chronic immune activation in the early stages of chronic conditions, consistent with Cluster 1, ultimately results in T cell exhaustion, dysregulated myeloid activation, and accelerated biological ageing that results in chronic disease progression, consistent with the clinical characteristics of Cluster 2.

A novel finding was the strong association between inflammageing, dysregulated haematopoiesis, platelet structure and function, and coagulopathy in Cluster 2. This points to wider systemic haematological derangement in this phenotype, and increased bleeding risk; a consideration in people with MLTC across many cardiovascular syndromes.

The association between the inflammageing phenotype, anaemia and transfusion provides a plausible explanation why interventions that reverse pre-surgery anaemia, or reduce transfusion, two clinical factors strongly associated with adverse outcome in surgery, do not reduce organ injury or improve prognosis.(47) Rather, these findings suggest that it is the unmeasured inflammageing phenotype that increases susceptibility to adverse outcome as observed in epidemiological studies, with anaemia and transfusion as bystander or subsequent events. These results lead us to hypothesise that immune system homeostasis presents a previously unrecognised target for pre-surgery organ protection interventions. This is supported by the causal inference analyses in UK Biobank which demonstrated that (genetic) modification of the molecular processes underlying the two MLTC phenotypes alters susceptibility to the stress of surgery measured as 90 day-mortality.

Finally, the very high levels of MODS (89%) in Cluster 2 notwithstanding, levels of MODS were also high (59%) in Cluster 1. This suggest that other disease mechanisms contribute to outcomes. In previous work, we have shown that MLTC are associated with myocardial changes consistent with biological ageing including altered RNA processing, DNA damage, nucleolar stress, senescence, and dysregulated energy metabolism versus non-MLTC.(48) Biological ageing processes increase susceptibility to cellular stress in experimental models. (49, 50) These results identify substantial baseline heterogeneity in cardiac surgery patients at baseline and suggest that a precision medicine approach to organ protection may be required to deliver clinical benefits.

Strengths and limitations

The clustering variables were selected based on biomarker availability across the existing cohorts. Alternative clustering variables will have yielded different results. In mitigation, the clustering biomarkers reflect the range of common chronic conditions observed in cardiac surgery patients at baseline. The resulting clusters also demonstrated consistent changes in immune homeostasis measured using different techniques across three independent cohorts recruited in multiple UK centres, implying validity. Poor reproducibility across clustering methods or datasets are key limitations of any clustering analyses. In this study, the
differences in organ injury between Clusters in the COPTIC validation cohort were less evident than for the development cohort. In mitigation, only 3 outcomes had objective ascertainment in the original COPTIC study; bleeding, AKI, and transfusion, all of which showed significant differences between clusters. These differences, along with the significant differences in biomarkers of immunological ageing in the COPTIC clusters, consistent with the differences in FACS detected leucocyte activation in the development cohort, and snRNAseq in the ObCARD cohort suggests external reproducibility of the results of the development analyses. As the snRNAseq analysis was restricted to atrial tissue, we also validated our findings externally in UK Biobank, with consistent results.

Conclusions

People with MLTC presenting for cardiac surgery demonstrate heterogeneous immune phenotypes in myocardium. The presence of an immunosenesence phenotype was associated with worse outcomes. Genetic modification of these processes altered susceptibility to adverse outcomes including mortality post-surgery, suggesting that immune system homeostasis represents a previously unrecognised target for organ protection interventions.

Contributions: FL designed the clinical clustering analyses, acquired the data, performed the statistical analyses, and wrote the manuscript. ASA designed the snRNAseq analyses, acquired the UK Biobank data, and undertook the bulk of the bioinformatics analysis. SS, SL, and KT performed snRNAseq experiments and bioinformatics analyses, and helped draft the paper. MR, TC, LJD, HA, TC, AM, GA, and MZ obtained the requisite ethical approvals, performed the clinical trials, and provided the tissue and blood samples. GC and MJW supervised SS and KT in the snRNAseq experiments and analyses. MJW co-designed the study and coordinated and supervised all laboratory analyses. GJM designed the study, wrote the manuscript, and is the guarantor of the study. All co-authors have critically reviewed the manuscript and agree to its publication.

Funding: The study was supported by British Heart Foundation grants RG/13/6/29947, CH/12/1/29419, and AA18/3/34220, and the Leicester NIHR Biomedical Research Centre. MR is a NIHR Clinical Lecturer. TC is an NIHR Academic Clinical fellow. FL is currently an employee for GlaxoSmithKline, but was an employee of the University of Leicester when this research was undertaken, supported by the British Heart Foundation grant CH/12/1/29419.

Competing Interests: The authors declare no competing interests.

Acknowledgements: The authors are grateful to Hasmukh Patel and Julie Chamberlain at the University of Leicester who performed biomarker analyses on the COPTIC samples, and Ms Bryony Eagle Hemming who undertook biomarker assays used for the development cohort.
analyses. We also express our gratitude to Dr Clelia Peano at Humanitas University in Milan for her expert advice and supervision on the development of the snRNAseq methods in human myocardium.
Figure legends (as they are listed in the manuscript)

**Figure 1. Workflow diagram.** Study 1 included 277 participants from the Revatio® for the Prevention of AKI (REVAKI-2) trial, (16) Red Cell washing for the Prevention of Organ Injury after Cardiac Surgery Trial (REDWASH), (15) and the Observational Case Control Study to Identify the Role of MV and MV Derived Micro-RNA in Post CArdiac Surgery AKI (MaRACAS). (14) These trials used cohort enrichment tools to select participants with MLTC at increased risk of organ injury. (4, 13) Study 2 included 800 participants from the Coagulation and Platelet Laboratory Testing in Cardiac Surgery (COPTIC) study.(11) Study 3 included 20 participants from A Case Control Study to Identify the Role of Epigenetic Regulation of Genes Responsible for Energy Metabolism and Mitochondrial Function in the Obesity Paradox in Cardiac Surgery (Ob-Card).(10)

**Figure 2. K-means Clustering of MLTC phenotypes**

Eleven biomarkers associated with MLTC from three cardiac surgery studies (n=277) were selected. A. The correlation matrix plot revealed no highly correlated variables. Variables were log-transformed if skewed and scaled to standard normal (mean 0, SD 1) before entering into cluster analysis. B. K-means clustering was performed. Two were considered as the optimal number of clusters after considering 23 indices. For visualisation, the classification plot was projected onto the subspace by the first two principal components. The clusters were characterised by comparing organ injury outcomes and cluster centres between phenotypes. C. Clusters demonstrate different susceptibility to organ injury as defined using consensus definitions or as measured using serum biomarkers.

Clusters in the COPTIC external validation dataset demonstrate differences in D. Baseline biomarkers of biological ageing and immunological ageing, and E. Organ injury and significant bleeding. Reported adverse clinical events were used to ascertain outcomes in the COPTIC study with the exception of bleeding which was the specified primary endpoint, and AKI which was ascertained from the analysis of routinely collected laboratory data.

**Figure 3. Major cell types and sub-cell types detected from single nuclei transcripts in human myocardial biopsies.** A. Dimensional plot showing cells of 11 major cell types detected B. Heatmap showing differences in annotated cells using canonical markers. Colours for expression markers are binary, either high (blue) or low (grey) C. Dimensional plot showing cardiomyocyte subtypes. Top gene signatures were derived by criteria of 1.5x detection, detection in at least 40% of cells in a subtype vs all other subtypes and higher expression (except CMs_2 with only lower gene expression). D. Top expressed genes in cardiomyocytes subtypes. Genes displayed include the top 25 genes by scaled expression and percent detected for each subtype combined with literature markers. There were overlaps in top25 for CMs_1 and CMs_2 including FGF12, CHRM2, ADGRL2, TMEM178B, SLIT3 and FHL2. E. Differentiation markers by CM subtype suggested that CM_0 to 2 are dedifferentiating. F.
Relative abundance of cardiomyocytes subtypes by Patient Cluster. **G.** Dimensional plot showing cardiac immune sub cell types. Gene signatures were derived in a similar fashion to cardiomyocytes. **H.** Relative abundance of cardiac immune subtypes by patient Cluster.

**Figure 4.** Metascape summary showing pathway enrichment for upregulated and downregulated genes in **A.** Cardiomyocytes. **B.** Lymphoid Cells, and **C.** Myeloid Cells. Bar plots are summarised from Metascape, full pathway results on GO Biological Process Annotation are presented in sTable13. **F.** Changes in expression in lymphoid cells **G.** Summary of biochemical pathways enriched by downregulated genes in lymphoid cells. Full pathway results on GO Biological Process Annotation are presented in **sTable 12.**

**Figure 5.** Internal and External Validation of snRNseq data. **A.** RNA scope for CM_0 markers ALCAM, BCL2, and MYBPC3 in myocardial biopsies using RNAscope (top 2 panels) and combined RNAscope Immunofluorescence (bottom 2 panels) in Cluster 1 and 2. **B.** Gene regulatory network in the CMs_0 cardiomyocyte subtype. **C.** Comparison of active regulatory networks in cardiomyocyte subtype between patient-clusters **D.** Effects of genetically-predicted expression on 90-day mortality post-surgery in UK Biobank participants.
Table 1. Baseline clinical and operative characteristics in the Development Cohort

<table>
<thead>
<tr>
<th></th>
<th>All Participants</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>74 (68 - 78)</td>
<td>74 (68 - 78)</td>
<td>73 (66 - 78)</td>
<td>0.4599</td>
</tr>
<tr>
<td>Sex – male</td>
<td>210 (76%)</td>
<td>157 (87%)</td>
<td>53 (55%)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>BMI (kg.m⁻²)</td>
<td>30.6 (27.1 - 34.8)</td>
<td>30.9 (27.6 - 34.9)</td>
<td>30.1 (25.4 - 33.3)</td>
<td>0.0507</td>
</tr>
<tr>
<td><strong>Medical history</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian Cardiology Society (CCS) Class</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>asymptomatic</td>
<td>102 (38%)</td>
<td>65 (36%)</td>
<td>37 (40%)</td>
<td>0.2681</td>
</tr>
<tr>
<td>I</td>
<td>83 (31%)</td>
<td>54 (30%)</td>
<td>29 (32%)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>73 (27%)</td>
<td>54 (30%)</td>
<td>19 (21%)</td>
<td></td>
</tr>
<tr>
<td>III &amp; IV</td>
<td>14 (5%)</td>
<td>7 (4%)</td>
<td>7 (8%)</td>
<td></td>
</tr>
<tr>
<td>Obese (BMI &gt;32)</td>
<td>106 (38%)</td>
<td>75 (41%)</td>
<td>31 (32%)</td>
<td>0.1362</td>
</tr>
<tr>
<td>Renal disease</td>
<td>28 (10%)</td>
<td>18 (10%)</td>
<td>10 (11%)</td>
<td>0.8576</td>
</tr>
<tr>
<td>Anaemia (Hb&lt;120 g.dL⁻¹ F, Hb&lt;130 g.dL⁻¹ M)</td>
<td>145 (53%)</td>
<td>65 (37%)</td>
<td>80 (85%)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diabetes</td>
<td>127 (47%)</td>
<td>80 (45%)</td>
<td>47 (51%)</td>
<td>0.3594</td>
</tr>
<tr>
<td>Stroke/TIA</td>
<td>25 (9%)</td>
<td>17 (10%)</td>
<td>8 (9%)</td>
<td>0.7978</td>
</tr>
<tr>
<td>Angina (CCS class II, III, IV) /previous MI</td>
<td>118 (43%)</td>
<td>81 (45%)</td>
<td>37 (39%)</td>
<td>0.3199</td>
</tr>
<tr>
<td>Heart failure (NYHA class III, IV)</td>
<td>60 (22%)</td>
<td>36 (20%)</td>
<td>24 (26%)</td>
<td>0.2425</td>
</tr>
<tr>
<td>Condition</td>
<td>MARACAS</td>
<td>REDWASH</td>
<td>REVAKI2</td>
<td>P-value</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Hypertension</td>
<td>11 (4%)</td>
<td>7 (4%)</td>
<td>4 (4%)</td>
<td>0.999</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1 (0.4%)</td>
<td>1 (0.6%)</td>
<td>0 (0%)</td>
<td>0.999</td>
</tr>
<tr>
<td>Pulmonary hypertension</td>
<td>58 (21%)</td>
<td>38 (21%)</td>
<td>20 (21%)</td>
<td>0.975</td>
</tr>
<tr>
<td>Extracardiac arteriopathy</td>
<td>15 (6%)</td>
<td>7 (4%)</td>
<td>8 (9%)</td>
<td>0.1079</td>
</tr>
<tr>
<td>Chronic Obstructive Pulmonary Disease</td>
<td>19 (7%)</td>
<td>1 (0.6%)</td>
<td>10 (11%)</td>
<td>0.0789</td>
</tr>
<tr>
<td>Asthma</td>
<td>2 (1%)</td>
<td>2 (1%)</td>
<td>0 (0%)</td>
<td>0.5454</td>
</tr>
<tr>
<td>Neurological diseases</td>
<td>5 (2%)</td>
<td>3 (2%)</td>
<td>2 (2%)</td>
<td>0.999</td>
</tr>
<tr>
<td>Cancer</td>
<td>8 (3%)</td>
<td>5 (3%)</td>
<td>3 (3%)</td>
<td>0.999</td>
</tr>
<tr>
<td>Liver diseases</td>
<td>3 (1%)</td>
<td>1 (0.6%)</td>
<td>2 (2%)</td>
<td>0.2763</td>
</tr>
<tr>
<td>Arthritis</td>
<td>2 (1%)</td>
<td>1 (0.6%)</td>
<td>1 (1%)</td>
<td>0.999</td>
</tr>
<tr>
<td>Multiple (2+ of the above) Long Term Conditions</td>
<td>225 (81%)</td>
<td>141 (78%)</td>
<td>84 (88%)</td>
<td>0.0515</td>
</tr>
</tbody>
</table>

**Baseline and operative characteristics**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MARACAS</th>
<th>REDWASH</th>
<th>REVAKI2</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemoglobin concentration (g.dL(^{-1}))</td>
<td>127 (114 - 138)</td>
<td>133 (123 - 144)</td>
<td>113.5 (103.5 - 122.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Haematocrit (%)</td>
<td>37.3 (33.9 - 40.6)</td>
<td>39.0 (36.5 - 42.2)</td>
<td>33.9 (31.4 - 36.8)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Serum creatinine (µmol.l(^{-1}))</td>
<td>87 (72 - 105)</td>
<td>88 (73 - 103)</td>
<td>86.5 (69 - 110.5)</td>
<td>0.7540</td>
</tr>
<tr>
<td>PaO2/FiO2 ratio</td>
<td>457 (410 - 533)</td>
<td>457 (410 - 533)</td>
<td>457 (410 - 533)</td>
<td>0.4177</td>
</tr>
<tr>
<td>eGFR (ml/min/1.73 m(^{2}))</td>
<td>74 (60 - 94)</td>
<td>75 (63 - 93)</td>
<td>72 (52 - 95)</td>
<td>0.0826</td>
</tr>
<tr>
<td><strong>Trial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARACAS</td>
<td>95 (34%)</td>
<td>75 (41%)</td>
<td>20 (21%)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>REDWASH</td>
<td>56 (20%)</td>
<td>7 (4%)</td>
<td>49 (51%)</td>
<td></td>
</tr>
<tr>
<td>REVAKI2</td>
<td>126 (45%)</td>
<td>99 (55%)</td>
<td>27 (28%)</td>
<td></td>
</tr>
<tr>
<td><strong>Surgery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

20
<table>
<thead>
<tr>
<th>Procedure</th>
<th>Median (Q1-Q3)</th>
<th>n (Percentage)</th>
<th>n (Percentage)</th>
<th>n (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CABG only</td>
<td>90 (33%)</td>
<td>64 (36%)</td>
<td>26 (27%)</td>
<td>0.2595</td>
</tr>
<tr>
<td>Valve only</td>
<td>92 (33%)</td>
<td>60 (33%)</td>
<td>32 (33%)</td>
<td></td>
</tr>
<tr>
<td>both or other cardiac surgery</td>
<td>94 (34%)</td>
<td>56 (31%)</td>
<td>38 (40%)</td>
<td></td>
</tr>
<tr>
<td>Cardiopulmonary Bypass Time (min)</td>
<td>103 (81 - 127)</td>
<td>102 (80 - 123)</td>
<td>106 (83 - 139.5)</td>
<td>0.2541</td>
</tr>
</tbody>
</table>

Note: Baseline clinical and operative characteristics are summarised as median (Q1-Q3) for continuous variables and n(percentage) for categorical variables.

**Labels:** NYHA New York Heart Association, eGFR, estimated Glomerular Filtration Rate, CABG, Coronary Artery Bypass Grafts, PaO2 Partial pressure of oxygen in arterial blood in mmHg, FiO2, % inspired oxygen concentration as a ratio, BMI, Body Mass Index
References

1. Khunti K, Sathanapally H, Mountain P. Multiple long term conditions, multimorbidity, and co-morbidities: we should reconsider the terminology we use. BMJ. 2023;383:p2327.
global health research. The Academy of Medical Sciences: The Academy of Medical Sciences; 2018.


**Aims**

**Development**
- To identify clusters using machine learning in MLTC cohort

**Co-red**
- Revatio® for the Prevention of AKI (REVAKI-2) trial, Red Cell washing for the Prevention of Organ Injury after Cardiac Surgery Trial (REDWASH), An Observational Case Control Study to Identify the Role of MV and MV Derived Micro-RNA in Post Cardiac Surgery AKI (MaRACAS)

**N=277 with complete baseline phenotyping and outcome data**

**Clustering Variables:** age, BMI, platelets, bilirubin, pro-NT BNP, TnI, eGFR, IL6, IL8, Hb, Transferrin, Transferrin Saturation, Fe, Ferritin, SuPAR, Plasma NGAL

**K means Clustering**
- Cluster 1: Cardiorenal/Anaemia Low
- Cluster 2: Cardiorenal/Anaemia High

**Internal Validation:**
- Agglomerative Hierarchical Clustering

**Analysis**

**Validation 1**
- To Validate MLTC clusters using machine learning

**Co-red**
- Coagulation and platelet laboratory testing in cardiac surgery (COPTIC) N=2400

**Cluster membership prediction with Development Cohort using age, platelet, Hb, FER, TRFE, troponin, IL6 (95% sensitivity)**

**N=800 with complete baseline phenotyping data**

**Clustering variables:** age, BMI, platelets, bilirubin, pro-NT BNP, TnI, eGFR, IL6, IL8, Hb, Transferrin, Ferritin

**K means Clustering**
- Cluster 1: Inflammation
- Cluster 2: Immunosenescence

**Pathway Analysis**

**Causal Inference Analysis of candidate genes/ pathways in UK Biobank**

**Validation 2**
- To identify cellular/ transcriptional characteristics of MLTC clusters

**Co-red**
- A Case Control Study to Identify the Role of Epigenetic Regulation of Genes Responsible for Energy Metabolism and Mitochondrial Function in the Obesity Paradox in Cardiac Surgery (Ob-Card) N=170

**N=20 with high quality snRNAseq data in human myocardium**

**Cluster membership prediction with Development Cohort using age, platelet, Hb, FER, TRFE, troponin, IL6 (95% sensitivity)**

**Cluster 1: Inflammation**
- N=13

**Cluster 2: Immunosenescence**
- N=7

**Cellular annotation**
- Cluster 1: Inflammation (n=13)
- Cluster 2: Immunosenescence (n=7)

**Gene Regulation Analysis**

**Differential Gene Analysis**

**Pathway Analysis**

**Development**
- To identify clusters using machine learning in MLTC cohort

**Co-red**
- To identify cellular/ transcriptional characteristics of MLTC clusters

**N=800 with complete baseline phenotyping data**

**Clustering variables:** age, BMI, platelets, bilirubin, pro-NT BNP, TnI, eGFR, IL6, IL8, Hb, Transferrin, Ferritin

**K means Clustering**
- Cluster 1: Inflammation
- Cluster 2: Immunosenescence

**Pathway Analysis**

**Gene Regulation Analysis**

**Causal Inference Analysis of candidate genes/ pathways in UK Biobank**

**K means Clustering**
- Cluster 1: Inflammation
- Cluster 2: Immunosenescence
Acute kidney injury (Any) Stage 2 or 3 AKI Myocardial injury Red cell transfusion Acute lung injury SIRS MODS

A. Heatmap of cluster centres in the COPTIC validation cohort

B. Organ Injury by K-Means Cluster in the Development Cohort

C. Organ Injury by K-Means Cluster in the Development Cohort

D. Heatmap of cluster centres in the COPTIC validation cohort

E. Organ Injury by K-Means Cluster in the COPTIC Cohort
CMs_3 signature (369): Cardiac contraction, translation, metabolic/mitochondrial activity & apoptosis

CMs_0 signature (4): immune response & cell apoptosis & polarisation

Macrophage Cell Clusters: Macrophage_0(AKTIPhi); MACs_1; MACs_2

T-Cells_0 signature (4): CD8+ T cells, CD4+CD8+ T cells

Cardiomyocyte cell clusters: CM_0: ALCAMhi, MKRN20Slo, FRMD4Ahi, and BCL2hi; CM_1: GALNT18hi, ENO4hi, CM_2: C8orf37-AS1hi, ZNF331hi, ZNF564hi, SLC45Alo; CM_3: DEShi, TCAPhi, ACTC1hi, MYH7hi
Differentially Expressed Genes

Summary of Pathway Enrichment for Downregulated Genes

Summary of Pathway Enrichment for Upregulated Genes

Cardiomyocytes

DEGs: 2776; Mostly downregulated

Lymphoid Cells

DEGs: 1115; Mostly downregulated

Myeloid Cells

DEGs: 3942; Mostly downregulated
**A.** Cluster 1 and Cluster 2

**B.** Regulatory Networks in genes differing between patient clusters in cardiomyocytes subset

**C.** Upregulated genes:
- ZMIZ1
- FOXP2
- TEAD1
- NFATC3

**D.** Downregulated genes:
- MYL4
- MYBPC3

**Transcription Factors**
- SPTBN1
- S100A9
- RPS2
- RYR2
- PPP1R9A
- HLA-B
- GMDS-DR

**Gene Expression QTL**
- log(p) values: 3.5, 4.0, 4.5, 5.0, 5.5
- **post surgery**

---

**Cluster 1 and Cluster 2**

**BCL2 and MYL4**

**ALCAM and MYL4**

**BCL2 and MYBPC3**

**ALCAM and MYBPC3**

---

[CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0)