AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Shared and distinct aspects of the sepsis transcriptomic response to fecal peritonitis and pneumonia

Authors: Katie L. Burnham[1], Emma E. Davenport[1], Jayachandran Radhakrishnan[1], Peter Humburg[1], Anthony C. Gordon[2], Paula Hutton[3], Eduardo Svoren-Jabalera[4], Christopher Garrard[3], Adrian V.S. Hill[1], Charles J. Hinds[5], Julian C. Knight[1*]

Affiliations:

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford.

[2]Section of Anaesthetics, Pain Medicine and Intensive Care, Imperial College, London.

[3]Adult Intensive Care Unit, John Radcliffe Hospital, Oxford.

[4]Queen Mary University of London, Queen's Hospital University Hospital Trust, London.

[5]William Harvey Research Institute, Barts and The London School of Medicine, Queen Mary University, London.

Corresponding author: Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN United Kingdom. Email: julian@well.ox.ac.uk, Telephone: +44 1865 287651

Author contributions: JCK and CJH conceived the study. KLB, EED, JR, ACG, JCK and CJH contributed to the study concept and design. KLB, EED, JR and ES performed the laboratory work. All authors participated in the acquisition, analysis, or interpretation of data. KLB, EED, JR, PeH and JCK contributed to the statistical analysis. PaH and AVSH provided administrative,

technical, or material support. JCK, KLB, EED and CJH contributed to drafting of the paper. JCK and CJH were responsible for obtaining funding. JCK led the study. All authors revised and approved the final version of the paper.

Funding: Supported by the National Institute for Health Research through the Comprehensive Clinical Research Network for patient recruitment; Wellcome Trust (Grants 074318 [JCK], and 090532/Z/09/Z [core facilities Wellcome Trust Centre for Human Genetics including High-Throughput Genomics Group]); European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement no. 281824 (JCK), the Medical Research Council (98082 [JCK]); UK Intensive Care Society; NIHR Oxford Biomedical Research Centre. AVSH is supported by a Wellcome Trust Senior Investigator Award (HCUZZ0) and ACG by an NIHR Clinician Scientist Fellowship. The organizations funding this study had no role in the design or conduct of the study; in the collection, management, analysis, or interpretation of the data; or in the preparation, review, or approval of the manuscript.

Conflict of interest: ACG reports personal fees and non-financial support from Orion Pharmaceuticals, grants and other from Tenax Therapeutics, other from Ferring Pharmaceuticals, personal fees from Baxter Healthcare, grants from HCA International, outside the submitted work. CJH reports grants from SIRIUS Genomics, outside the submitted work. The other authors declare no competing interests.

Running head: Sepsis response transcriptomics

Descriptor: 4.12 Sepsis/Multiple Organ Failure

Word count: 3442

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

At a Glance Commentary

Scientific Knowledge on the Subject

There is significant heterogeneity in the septic response, which has hindered efforts to understand pathophysiology and develop targeted therapies. Molecular approaches may provide insights into variation in the host response, enabling biomarker development. Recent evidence suggests transcriptomic sepsis response signatures (SRS) can define patient subgroups associated with early outcome in sepsis due to community acquired pneumonia (CAP).

What This Study Adds to the Field

This study provides the first substantive analysis of the transcriptomic response in patients admitted to intensive care with sepsis due to fecal peritonitis (FP). Comparison with sepsis due to CAP and with non-septic controls demonstrates a shared sepsis response, independent of source of infection, which involves a significant proportion of the transcribed genome and overlaps with the "genomic storm" following trauma. We find evidence of SRS groups in FP patients predictive of early mortality, with group membership changing over time in some patients. We show that the major predictor of variation in gene expression between sepsis patients is SRS group rather than source of infection, with only a small number of genes differentially regulated according to the latter, enriched for interferon signaling and antigen presentation. These findings highlight opportunities for patient stratification in sepsis.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

Abstract

Rationale: Heterogeneity in the septic response has hindered efforts to understand pathophysiology and develop targeted therapies. Source of infection, with different causative organisms and temporal changes, might influence this heterogeneity.

Objectives: To investigate individual and temporal variation in the transcriptomic response to sepsis due to fecal peritonitis, and to compare with community acquired pneumonia.

Methods: We performed genome-wide gene expression profiling in peripheral blood leukocytes for adult patients admitted to intensive care with sepsis due to fecal peritonitis (n=117) or community acquired pneumonia (n=126), and non-septic controls (n=10).

Measurements and Main Results: A substantial portion of the transcribed genome (18%) was differentially expressed compared to controls, independent of source of infection, with EIF2 signaling the most enriched canonical pathway. We identify two sepsis response signature subgroups in fecal peritonitis associated with early mortality (p-value=0.01, hazard ratio=4.78). We define gene sets predictive of SRS group, and serial sampling demonstrates subgroup membership is dynamic during ICU admission. We find SRS is the major predictor of transcriptomic variation; a small number of genes (n=263) were differentially regulated according to the source of infection, enriched for interferon signaling and antigen presentation. We define temporal changes in gene expression from disease onset involving phagosome formation, NK cell and IL-3 signaling.

Conclusions: The majority of the sepsis transcriptomic response is independent of source of infection and includes signatures reflecting immune response state and prognosis. A modest number of genes show evidence of specificity. Our findings highlight opportunities for patient stratification and precision medicine in sepsis.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Word count: 249

Key words: gene expression, septic, immune, patient stratification

Introduction

Sepsis is the life-threatening organ dysfunction caused by a dysregulated host response to infection (1). Therapeutic options remain limited despite extensive efforts to develop and refine treatment strategies (2). New insights into pathophysiology and the development of targeted treatments appropriate for individual patients at specific stages in the illness are urgently required (3). To be successful, this approach requires clearer understanding of heterogeneity in the sepsis response, in which pro-inflammatory and immunosuppressed states are dynamic and frequently co-exist (4). We recently identified distinct transcriptomic sepsis response signatures (SRS) in peripheral blood leukocytes from patients with community acquired pneumonia (CAP), informative for immune response states and outcome (5). In particular, patients with SRS1 exhibit an immunosuppressed phenotype associated with higher early mortality, with features of endotoxin tolerance, T-cell exhaustion and downregulation of HLA class II.

It is not yet known whether comparable SRS are present in sepsis caused by other sources of infection. Additionally, the extent to which the source of infection contributes to heterogeneity in the transcriptomic response is unclear, although there is evidence that gene expression signatures can distinguish between Gram positive, Gram negative and viral etiologies, and may be useful in the diagnosis of CAP (6-13). In fecal peritonitis (FP), sepsis is triggered by a poly-microbial infection within the peritoneal cavity, complicated by the release of damage associated molecular patterns (DAMPS) and the effects of anaesthesia (14). Conversely, CAP is caused by specific bacterial or atypical pathogens, and is commonly preceded or caused by viral infection (15). Antibiotic choice, for example macrolides for CAP, may also have immunomodulatory effects (16, 17). Finally, while FP patients usually have a rapid onset of illness, CAP patients are often unwell for many days prior to ICU admission.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

FP is a common cause of sepsis with a high mortality (18), in which longitudinal analysis is tractable given that a defined time of onset can be estimated. We hypothesized that sepsis response signatures similar to those seen in CAP would be present in FP patients but that aspects of the transcriptomic host response would be dependent on the source of infection and stage of illness. Here, we investigate how patterns of gene expression in leukocytes are influenced by source of infection for FP and CAP, and how they vary between individual patients and within patients over time.

Some of the results of these studies have been previously reported in the form of abstracts (19, 20).

Methods

Study design and participants

The objective was to characterize the transcriptomic response to sepsis caused by FP, including an analysis of temporal changes and a comparison with sepsis caused by CAP. The subjects were adult patients admitted to ICU with sepsis as part of the UK GAinS study (www.ukccg-gains.org) with pre-defined inclusion and exclusion criteria (described in online data supplement). FP was diagnosed at laparotomy as inflammation of the peritoneal membrane secondary to large bowel perforation and fecal contamination (18). CAP was defined as a febrile illness associated with cough, sputum production, breathlessness, leukocytosis and radiological features of pneumonia, acquired in the community or within two days of ICU admission (21, 22). This was a prospective observational study. The transcriptomic response was investigated in peripheral blood leukocytes and compared between patients and over time, and with controls

undergoing elective cardiac surgery (described in online data supplement). Discovery and validation cohorts were used to identify shared and specific gene expression patterns among FP and CAP patients (Supplementary Fig E1).

Sample Collection

Blood samples (5ml) were collected (Vacuette EDTA tubes) on the first, third, and/or fifth day after ICU admission. The total blood leukocyte population was isolated using LeukoLOCK filters (Ambion), stabilized using RNA*later,* and total RNA extracted (described in online data supplement). Extensive, anonymized clinical information was recorded using an electronic Case Report Form (eCRF) (5).

Microarray analysis

Genome-wide gene expression was quantified using Illumina Human-HT-12 v4 Expression BeadChips (47,231 probes) with sample processing, data preparation, background subtraction, transformation and normalization using the vsn package (23) (described in the online data supplement). Gene expression data is available through ArrayExpress (accession: E-MTAB-5273/E-MTAB-5274).

Statistical analysis

Analysis was carried out using R unless otherwise specified, with statistical tests, power calculations, differential gene expression, enrichment testing, predictive modelling, and cluster analysis described in the online data supplement.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Results

Transcriptomic response to fecal peritonitis: identification of sepsis response signatures

We quantified genome-wide gene expression for 117 FP and 126 CAP adult patients admitted to ICU with sepsis, analyzing RNA from blood leukocytes rapidly isolated at the bedside. All patients showed evidence of organ dysfunction based on SOFA scores during ICU admission. These patients were recruited concurrently and processed in parallel as discovery (64 FP, 73 CAP) and validation cohorts (53 FP, 53 CAP) (these 53 CAP cases were previously published (5)). We also analyzed samples obtained from ten subjects prior to elective cardiac surgery as non-septic controls in the discovery cohort. Demographics and clinical covariates for the discovery set (FP n=64, CAP n=73 patients; 221 samples, Supplementary Fig. E1) are shown (Table 1, Supplementary Table E1A).

We investigated whether sepsis response subgroups based on global patterns of gene expression were present in FP. The combined FP cohort (147 samples from 117 patients) had a 28-day mortality of 13%; the mean age was 66 years and 50.4% of patients were male (Table 2, Supplementary Table E1B). We applied two complementary approaches to this FP cohort: (1) hypothesis-free unsupervised hierarchical clustering based on observed patterns of variation in gene expression and (2) assignment based on expression of a specific seven gene set previously shown to be predictive of SRS group membership in CAP (5).

We first analyzed variation in global gene expression (2716 probes, 10% most variably expressed genes) using agglomerative hierarchical clustering followed by consolidation of group

membership using k-means. This approach revealed two distinct patient groups, SRS1_FP (n=68, 46%) and SRS2_FP (n=79, 54%) (Fig.1 A, B). Following group assignment, 1075 genes were found to be differentially expressed between these groups (Fig. 1C, Supplementary Table E1C), showing strong correlation with those previously observed in CAP patients (5) (Pearson's r=0.791; Supplementary Fig. E2). To determine the functional overlap with SRS groups identified in CAP, we performed pathway, function enrichment, and upstream regulator analysis (Fig. 1D, Supplementary Table E1D). Cell death, apoptosis, necrosis, T cell activation, and endotoxin tolerance were significantly enriched biological functions in SRS1 for both sources of infection. We proceeded to test for association with outcome in FP, and found that SRS1_FP group membership was associated with higher early mortality (14 day mortality log rank test P=0.0096, 18.8% vs 4.3%, HR=4.78 (1.29-17.65)) (Fig. 1E). The effect of SRS group membership remained after inclusion of age, SOFA score, Charlson Comorbidity Index, and sex in a cox proportional hazards model (p=0.037, HR=4.23 (1.09-16.39)).

To further validate our findings, we adopted a second approach, namely assignment of SRS group membership in this FP cohort using the seven gene set (*DYRK2*, *CCNB1IP1*, *TDRD9*, *ZAP70*, *ARL14EP*, *MDC1* and *ADGRE3*) we previously established in CAP patients (5) (Supplementary Fig. E3). This showed strong concordance with the results obtained by unsupervised analysis (misclassification rate 21.1%, AUC 0.923; Fig. 1F). The groups defined using the seven gene set showed a significant difference in early mortality rates (log rank test P=0.030 FP patients), a differential gene expression signature strongly correlated with that seen in the original CAP analysis (r=0.845), and pathway enrichment comparable to findings in CAP patients. We also derived a six gene set (*CD163, ZDHHC19, MME, FAM89A, ZBP1, B3GNT2)*

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

from the FP data predictive of SRS_FP group membership, with a 4.1% misclassification rate based on internal leave-one-out cross-validation (AUC 0.975; Fig. 1F). When applied to the original CAP cohort (5), this six gene set again performed well (misclassification rate 27.9%, AUC 0.931).

A poor outcome subgroup has previously been reported in children with sepsis (24). We investigated the similarity between pediatric endotype A and SRS1 by comparative differential gene expression analysis. Although there was some overlap in pathway enrichment (e.g. T and B cell receptor signaling), the gene expression patterns that distinguish SRS groups were not enriched in the pediatric endotype contrast (Supplementary Fig. E4).

Temporal changes in sepsis response signatures

Serial sampling on day 1, day 3, and day 5 following ICU admission provides the opportunity to investigate the relationship between SRS group membership and disease progression. We found that 11 out of 24 FP patients (46%) with serial samples moved between groups over time; 10 from SRS1 to SRS2, of whom only 1 died (Fig. 1G). Thirteen patients (54%) remained in the same SRS group; 3/7 patients remaining in SRS1 died, compared to no deaths amongst the 6 remaining in SRS2. This movement between SRS groups involves large changes in gene expression, illustrated by *CD163,* encoding an innate immune sensor for bacteria, and one of the six gene classifiers for SRS_FP (Fig. 1H, Supplementary Fig. E5).

Influence of source of infection on the transcriptomic response

We proceeded to further characterize and compare gene expression in FP and CAP patients to determine the relative importance of shared and specific features of the transcriptomic response. As expected, in the discovery cohort GI-related comorbidities were more common in FP patients, while CAP patients had more respiratory comorbidities, higher respiratory rate and oxygenation requirements, and a higher lymphocyte count (Table 1 and Supplementary Table E1A). Although mortality did not differ significantly, CAP patients had higher APACHE and SOFA scores (Table 1).

To understand the relationships between source of infection, SRS group, and heterogeneity in the septic response we performed principal component analysis using the top 10% most variably expressed transcripts. Non-septic controls cluster together and are clearly distinct from septic patients (Fig. 2A). Among the combined population of septic patients, there was clear segregation based on SRS group but not source of infection (Fig. 2A).

In order to further elucidate the drivers of variation in the septic response, we calculated correlations between the observed variance in global gene expression for the sepsis samples (the first six principal components of variance, representing ~50% of the variance in the data), a comprehensive set of clinical covariates, and SRS group (Fig. 2B, Supplementary Table E1E). This showed variation in gene expression was most strongly correlated with SRS group membership (PC1: r=0.77, FDR=$<2.2\text{x}10^{-16}$). More modest correlation was seen with source of infection (PC1: r=-0.417, FDR=$6.12\text{x}10^{-9}$), bicarbonate levels (PC1: r=0.35, FDR=$2.2\text{x}10^{-8}$) and

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

neutrophil count (PC1: r=-0.29, FDR=2.0x10^{-3}). When the dataset was divided into CAP and FP samples and the analysis repeated, we found SRS remained the most significant correlate (Supplementary Fig. E6).

When we compared all patients in the discovery cohort (FP n=64, CAP n=73) with non-septic controls (n=10) we found that sepsis was associated with differential regulation of a large proportion of the transcribed genome (18.0% of transcripts assayed, 4881 probes; >1.5 fold change (FC), FDR <0.05; Supplementary Table E1F). Pathway analysis showed EIF2 signaling was the top canonical pathway, consistent with the reported role for this translational initiation factor in response to viral and bacterial infection (25), and predicted key upstream regulators IL13, ATB1, TGFB1 and IL2 (Supplementary Table E1G). Using the same methodology, we contrasted global gene expression differences in sepsis compared to controls with those observed in the previously published genomic response to trauma versus healthy subjects (26) within the same time frame as the sepsis cohort (5 days following injury). We found that of the genes differentially expressed in the sepsis response and measured in the trauma cohort, the majority (n=1884, 64%) were also differentially expressed (Fig. 2C). Commonality was seen with EIF2 signaling and inflammation-related pathways enriched in both sepsis and trauma responses, while the inflammatory response to organismal injury was specific to trauma, and TNFR1 signaling to sepsis (Fig. 2C, Supplementary Table E1H).

When we compared CAP vs controls and FP vs controls, gene expression patterns were highly correlated, demonstrating that most sepsis response pathways are common and independent of source of infection (Fig. 2D, 3A, B; Supplementary Table E1I, J).

To further investigate differential gene expression between patients with sepsis due to FP or CAP, we compared the first available samples for each patient following ICU admission (FP n=64, CAP n=73). We found 310 probes (263 genes) differentially expressed between the two sources of infection (FC >1.5, FDR <0.05; Fig. 3C, Supplementary Table E1K), significantly more than expected by chance based on permutation analysis (Supplementary Fig. E7). We noted significant differences in total and differential cell counts between CAP and FP. The inclusion of differential cell counts in the regression model had minimal influence, with a similar list of differentially regulated genes and fold changes strongly correlated with the original analysis (Supplementary Fig. E8, Supplementary Table E1K). IFN signaling (P $6.1x10^{-10}$) and antigen presentation (P $1.6x10^{-8}$) were the most significant enriched canonical pathways, upregulated in CAP patients compared to FP (Fig. 3D, Supplementary Table E1L). Enriched networks involved IFNα/β and the antimicrobial/inflammatory response (P $1x10^{-39}$) (Fig. 3E). *IFIT1*, *IFIT2*, and *IFIT3* (interferon-induced antiviral genes), *BTN3A3* (MHC class I gene), *IFIH1* (sensor of viral nucleic acids) and *OAS2* (interferon-induced antiviral enzyme) were all up-regulated in CAP patients. IFNs, endotoxin and TNF were found to be significant upstream regulators (Fig. 3D, Supplementary Table E1L).

We confirmed our findings in a validation set comprising 53 FP and 53 CAP patients prospectively recruited to the UK GAinS study (Supplementary Fig. E1). We found SRS group, source of infection, cell count, and day of sampling were strongly associated with the first six principal components of gene expression; genes differentially expressed between FP/CAP correlated with those found in the discovery cohort (Pearson's r=0.83, P < $2.2x10^{-16}$;

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Supplementary Table E1M). Pathway analysis demonstrated high concordance of enriched pathways and functions (Fig. 3D, Supplementary Table E1N). We found a gene set comprising *EPHB1*, *NQO2*, *ARG1*, *HMGB2, FGL2,* and *GPR162* was predictive of source of infection (FP or CAP) for the discovery cohort. We applied a sparse regression method to show that for the discovery cohort the misclassification rate was 29.9% with internal leave-one-out validation (AUC 0.760); for the validation cohort the test error was 28.3% (AUC 0.798; Supplementary Fig E9). There was no difference in the proportion of different sampling days following ICU admission between FP and CAP samples in the analyzed cohorts (Supplementary Fig. E10, chi-square P=0.350).

Given a prominent viral signature amongst genes differentially expressed between FP and CAP patients, we investigated if this came from a subset of CAP patients with a viral infection. Only 6/73 CAP patients in the discovery cohort had a confirmed viral infection; we therefore analyzed the transcriptomic differences between patients with and without confirmed viral infection in our larger previously published CAP cohort (25 vs 240 patients) (5). We identified 39 differentially expressed genes (FDR <0.05, FC >1.5) (Fig. 3F, Table E1O) including *IFI27*, an interferon alpha-inducible protein reported as a marker of influenza infection (27); and *LAMP3*, a dendritic cell glycoprotein induced by influenza A infection (28). Pathway analysis showed most significant enrichment for pattern recognition receptors, IFN signaling and IRFs (Supplementary Table E1P). We then compared this gene set to the genes differentially expressed between FP and CAP patients and found significant enrichment of 24/39 genes ($P < 1x10^{-4}$).

Temporal changes in gene expression

Interpretation of gene expression patterns in sepsis may be complicated by dynamic differences in the host response over time (10, 29). We explored temporal changes in gene expression for FP and CAP patients in whom serial samples following ICU admission were available. We first sought to identify genes that varied over time using a repeated measures regression model. We found 714 genes were significantly differentially expressed in FP sepsis patients between days 1 and 5, compared to only 80 genes in CAP patients (FDR <0.05, FC >1.5) (Fig. 4A, Supplementary Table E1Q). We further analyzed our data using a multivariate empirical Bayes' model, which ranks genes based on differential expression for longitudinal series involving multiple biological conditions (30), restricting this analysis to patients where samples were available at all 3 time points. Notably, temporal changes in gene expression were again more pronounced in FP patients (Supplementary Table E1R), and the specific genes whose expression changed over time differed between FP and CAP (Fig. 4B,C).

In FP patients it was possible to estimate with reasonable accuracy the time of onset of the acute event. We calculated days from disease onset for each FP sample and used this to investigate temporal changes. We found that using a linear model with limma analysis, 140 genes show significant changes in expression over time (FDR <0.05) (Supplementary Table E1S) including *AAK1* and *SNN* (Fig. 4D). Pathway analysis was significant for genes involved in phagosome formation, NK cell and IL-3 signaling with evidence of enrichment for specific biological functions notably cellular degranulation, chemotaxis, leukocyte activation, macrophage adhesion, phagocytosis and bacterial infection (Supplementary Table E1T).

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Discussion

We have characterised the transcriptomic response to sepsis caused by FP and found evidence of sepsis response signatures associated with outcome, patients in SRS1 having higher early mortality. Transcriptomic features of endotoxin tolerance and pathway enrichment for cell death, apoptosis, necrosis and T cell exhaustion are consistent with animal models and human studies demonstrating the importance of immune compromise in sepsis pathogenesis and as a determinant of poor outcomes (4, 31, 32). The FP dataset allowed us to explore evolution of SRS membership. We find a significant number of patients (46%) switch SRS group in the first 5 days of ICU admission, the majority moving from SRS1 to SRS2. Persistence of SRS1 is associated with a poor outcome whereas maintenance or recovery of immune competence (SRS2) is associated with survival. These findings further support the concept that SRS group membership reflects clinically important biological differences (5) and suggest that if used as a biomarker, the transcriptomic response signature should be determined at the time a therapeutic intervention is being considered. Establishing the immune response state of a patient could enable individually tailored immunotherapy and monitoring the response to treatment.

We found that the transcriptomic response is, to a large extent, shared between the two sources of infection analysed here, with the expression of only a modest number of genes being dependent on source of infection. This shared sepsis response involves a significant proportion of the transcribed genome and overlapped with the "genomic storm" following trauma although we observed some differences for example involving TNFR1 signaling (26). Comparing gene expression between FP and CAP, the most enriched networks involved IFNα/β and the antimicrobial/inflammatory response. These differences seem to be predominantly driven by

viral respiratory infection within the CAP cohort, although given current difficulties in pathogen detection the biological and clinical interpretation of such transcriptomic differences remains challenging. While some previous studies have supported a common transcriptional septic response independent of pathogen (33), others have reported that expression signatures can discriminate between infecting organisms (6-8, 13) although these findings remain controversial (9, 10). For patients admitted to ICU with suspected CAP, a 78 transcript signature has been reported to differentiate cases of CAP from non-CAP with the *FAIM3*:*PLAC8* gene expression ratio proposed as a diagnostic biomarker (11).

We found temporal changes were more pronounced in FP, with more than eight times as many genes differentially expressed between admission and day 5 than in CAP patients, and involved phagosome formation, NK cell signaling, IL-3 signaling, leukocyte activation, mitochondrial damage, and apoptosis. To date, time series analyzing changes in gene expression have focused on animal models of sepsis (34), the endotoxin response in healthy volunteers (35), the response to trauma and subsequent infection (26), and early events in sepsis (29), with a recent analysis of time-matched cohorts defining a gene set distinguishing sterile inflammation from infectious inflammation (12). In general, clearer resolution of temporal differences will be critical to resolving heterogeneity in observed sepsis responses within and between patients.

Future work should include comprehensive pathogen phenotyping, for example using metagenomic sequencing, given the recognized high proportion of culture-negative patients in sepsis (36, 37). Comparison of FP patients with specific controls, such as a group of patients undergoing laparotomy for non-infectious indications would control for the effects of DAMP-

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

mediated signaling from the surgical procedure, and any modulating influences on the transcriptome from general anesthesia, while inclusion of patients undergoing laparotomy for gastric or small-bowel perforation would to some extent control for different pathogens. Differences in differential cell count are potential confounders in transcriptomic analysis (10) and, although the results presented were robust to differential leukocyte count, cell type-specific transcriptomic profiling will be important for future studies (38).

Defining the most robust and informative predictive gene set for SRS membership based on transcriptomics will require prospective large-scale validation using an appropriate point of care test. We note that the current seven gene set established in CAP patients and validated in FP is consistent with a key role for a dysfunctional immune response in sepsis. This predictive gene set includes genes involved in cell growth (*DYRK2*), cell cycle (*CCNB1IP1*), stem cell maintenance (*TDRD9*) and DNA damage (*ADGRE3*), consistent with the observed pathway enrichment (cell death, apoptosis, necrosis); and with immune response through lymphocyte activation (*ZAP70*), MHC class II export (*ARL14EP*) and myeloid cell interactions in immunity (*MDC1*). The lack of overlap with paediatric sepsis endotypes we observed requires further validation and may reflect developmental differences impacting on pathophysiology and host immune dysfunction (39, 40).

We have addressed sepsis heterogeneity using a transcriptomic approach in patients with FP and CAP, demonstrating a shared sepsis response with distinct SRS groups that are dynamic, reflect the underlying biological response, and are informative for prognosis. Our findings also provide some evidence for differential patterns of expression between CAP and FP, two of the most common causes of sepsis. Our analysis, coupled with recent advances in the field (3, 4, 11, 41),

highlights the opportunity to develop novel therapeutic interventions that can be targeted and appropriately timed to individual patients based on transcriptomic signatures, providing opportunities for precision medicine in sepsis.

Acknowledgments: We thank all participating patients, GAinS investigators, and recruiting hospitals (Supplementary Table E2).

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, Hotchkiss RS, Levy MM, Marshall JC, Martin GS, Opal SM, Rubenfeld GD, van der Poll T, Vincent JL, Angus DC. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016; 315: 801-810.
2. Marshall JC. Why have clinical trials in sepsis failed? *Trends Mol Med* 2014; 20: 195-203.
3. Cohen J, Vincent JL, Adhikari NK, Machado FR, Angus DC, Calandra T, Jaton K, Giulieri S, Delaloye J, Opal S, Tracey K, van der Poll T, Pelfrene E. Sepsis: a roadmap for future research. *Lancet Infect Dis* 2015; 15: 581-614.
4. Hotchkiss RS, Monneret G, Payen D. Sepsis-induced immunosuppression: from cellular dysfunctions to immunotherapy. *Nat Rev Immunol* 2013; 13: 862-874.
5. Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, Rautanen A, Gordon AC, Garrard C, Hill AVS, Hinds CJ, Knight JC. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir Med* 2016; 4: 259-271.
6. Ramilo O, Allman W, Chung W, Mejias A, Ardura M, Glaser C, Wittkowski KM, Piqueras B, Banchereau J, Palucka AK, Chaussabel D. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 2007; 109: 2066-2077.
7. Yu SL, Chen HW, Yang PC, Peck K, Tsai MH, Chen JJ, Lin FY. Differential gene expression in gram-negative and gram-positive sepsis. *Am J Respir Crit Care Med* 2004; 169: 1135-1143.
8. Feezor RJ, Oberholzer C, Baker HV, Novick D, Rubinstein M, Moldawer LL, Pribble J, Souza S, Dinarello CA, Ertel W, Oberholzer A. Molecular characterization of the acute inflammatory response to infections with gram-negative versus gram-positive bacteria. *Infect Immun* 2003; 71: 5803-5813.
9. Tang BM, McLean AS, Dawes IW, Huang SJ, Cowley MJ, Lin RC. Gene-expression profiling of gram-positive and gram-negative sepsis in critically ill patients. *Crit Care Med* 2008; 36: 1125-1128.
10. Maslove DM, Wong HR. Gene expression profiling in sepsis: timing, tissue, and translational considerations. *Trends Mol Med* 2014; 20: 204-213.
11. Scicluna BP, Klein Klouwenberg PM, van Vught LA, Wiewel MA, Ong DS, Zwinderman AH, Franitza M, Toliat MR, Nurnberg P, Hoogendijk AJ, Horn J, Cremer OL, Schultz MJ, Bonten MJ, van der Poll T. A molecular biomarker to diagnose community-acquired pneumonia on intensive care unit admission. *Am J Respir Crit Care Med* 2015; 192: 826-835.
12. Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Science Transl Med* 2015; 7: 287ra271.
13. Sweeney TE, Wong HR, Khatri P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. *Science Transl Med* 2016; 8: 346ra391.
14. Krepel CJ, Gohr CM, Edmiston CE, Condon RE. Surgical sepsis: constancy of antibiotic susceptibility of causative organisms. *Surgery* 1995; 117: 505-509.

15. Marik PE. The clinical features of severe community-acquired pneumonia presenting as septic shock. Norasept II Study Investigators. *J Crit Care* 2000; 15: 85-90.
16. Weiss T, Shalit I, Blau H, Werber S, Halperin D, Levitov A, Fabian I. Anti-inflammatory effects of moxifloxacin on activated human monocytic cells: inhibition of NF-kappaB and mitogen-activated protein kinase activation and of synthesis of proinflammatory cytokines. *Antimicrob Agents Chemother* 2004; 48: 1974-1982.
17. Kovaleva A, Remmelts HH, Rijkers GT, Hoepelman AI, Biesma DH, Oosterheert JJ. Immunomodulatory effects of macrolides during community-acquired pneumonia: a literature review. *J Antimicrob Chemother* 2012; 67: 530-540.
18. Tridente A, Clarke GM, Walden A, McKechnie S, Hutton P, Mills GH, Gordon AC, Holloway PA, Chiche JD, Bion J, Stuber F, Garrard C, Hinds CJ, Gen OI. Patients with faecal peritonitis admitted to European intensive care units: an epidemiological survey of the GenOSept cohort. *Intensive Care Med* 2014; 40: 202-210.
19. Burnham K, Davenport E, Radhakrishnan J, Svoren-Jabalera E, Humburg P, Hutton P, Gordon A, Hill A, Hinds C, Knight J. Shared and specific features of the individual transcriptomic response in sepsis due to faecal peritonitis. Poster presented at the American Society of Human Genetics, Vancouver Canada. 2016.
20. Burnham KL, Davenport EE, Radhakrishnan J, Humburg P, Hutton P, Garrard C, Hinds CJ, Knight JC. Heterogeneity in the individual transcriptomic response to severe sepsis. Talk presented at the 13th International Congress of Human Genetics, Kyoto Japan. 2016.
21. Angus DC, Marrie TJ, Obrosky DS, Clermont G, Dremsizov TT, Coley C, Fine MJ, Singer DE, Kapoor WN. Severe community-acquired pneumonia: use of intensive care services and evaluation of American and British Thoracic Society Diagnostic criteria. *Am J Respir Crit Care Med* 2002; 166: 717-723.
22. Walden AP, Clarke GM, McKechnie S, Hutton P, Gordon AC, Rello J, Chiche JD, Stueber F, Garrard CS, Hinds CJ. Patients with community acquired pneumonia admitted to European Intensive Care Units: an epidemiological survey of the GenOSept cohort. *Crit Care* 2014; 18: R58.
23. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 2002; 18 Suppl 1: S96-104.
24. Wong HR, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Willson DF, Freishtat RJ, Anas N, Meyer K, Checchia PA, Monaco M, Odom K, Shanley TP. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med* 2009; 7: 34.
25. Shrestha N, Bahnan W, Wiley DJ, Barber G, Fields KA, Schesser K. Eukaryotic initiation factor 2 (eIF2) signaling regulates proinflammatory cytokine expression and bacterial invasion. *J Biol Chem* 2012; 287: 28738-28744.
26. Xiao W, Mindrinos MN, Seok J, Cuschieri J, Cuenca AG, Gao H, Hayden DL, Hennessy L, Moore EE, Minei JP, Bankey PE, Johnson JL, Sperry J, Nathens AB, Billiar TR, West MA, Brownstein BH, Mason PH, Baker HV, Finnerty CC, Jeschke MG, Lopez MC, Klein MB, Gamelli RL, Gibran NS, Arnoldo B, Xu W, Zhang Y, Calvano SE, McDonald-Smith GP, Schoenfeld DA, Storey JD, Cobb JP, Warren HS, Moldawer LL, Herndon DN, Lowry SF, Maier RV, Davis RW, Tompkins RG, Inflammation, Host Response to Injury Large-Scale

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Collaborative Research P. A genomic storm in critically injured humans. *J Exp Med* 2011; 208: 2581-2590.
27. Ioannidis I, McNally B, Willette M, Peeples ME, Chaussabel D, Durbin JE, Ramilo O, Mejias A, Flano E. Plasticity and virus specificity of the airway epithelial cell immune response during respiratory virus infection. *J Virol* 2012; 86: 5422-5436.
28. Zhou Z, Xue Q, Wan Y, Yang Y, Wang J, Hung T. Lysosome-associated membrane glycoprotein 3 is involved in influenza A virus replication in human lung epithelial (A549) cells. *Virology Journal* 2011; 8: 384.
29. Cazalis MA, Lepape A, Venet F, Frager F, Mougin B, Vallin H, Paye M, Pachot A, Monneret G. Early and dynamic changes in gene expression in septic shock patients: a genome-wide approach. *Intensive Care Med Exp* 2014; 2: 20.
30. Tai YC, Speed TP. On gene ranking using replicated microarray time course data. *Biometrics* 2009; 65: 40-51.
31. Boomer JS, To K, Chang KC, Takasu O, Osborne DF, Walton AH, Bricker TL, Jarman SD, 2nd, Kreisel D, Krupnick AS, Srivastava A, Swanson PE, Green JM, Hotchkiss RS. Immunosuppression in patients who die of sepsis and multiple organ failure. *JAMA* 2011; 306: 2594-2605.
32. Pena OM, Hancock DG, Lyle NH, Linder A, Russell JA, Xia J, Fjell CD, Boyd JH, Hancock RE. An Endotoxin Tolerance Signature Predicts Sepsis and Organ Dysfunction at Initial Clinical Presentation. *EBioMedicine* 2014; 1: 64-71.
33. Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol* 2005; 3: 281-294.
34. Mira JC, Szpila BE, Nacionales DC, Lopez MC, Gentile LF, Mathias BJ, Vanzant EL, Ungaro R, Holden D, Rosenthal MD, Rincon J, Verdugo PT, Larson SD, Moore FA, Brakenridge SC, Mohr AM, Baker HV, Moldawer LL, Efron PA. Patterns of gene expression among murine models of hemorrhagic shock/trauma and sepsis. *Physiol Genomics* 2016; 48: 135-144.
35. Talwar S, Munson PJ, Barb J, Fiuza C, Cintron AP, Logun C, Tropea M, Khan S, Reda D, Shelhamer JH, Danner RL, Suffredini AF. Gene expression profiles of peripheral blood leukocytes after endotoxin challenge in humans. *Physiol Genomics* 2006; 25: 203-215.
36. Phua J, Ngerng W, See K, Tay C, Kiong T, Lim H, Chew M, Yip H, Tan A, Khalizah H, Capistrano R, Lee K, Mukhopadhyay A. Characteristics and outcomes of culture-negative versus culture-positive severe sepsis. *Crit Care* 2013; 17: R202.
37. Vincent JL, Rello J, Marshall J, Silva E, Anzueto A, Martin CD, Moreno R, Lipman J, Gomersall C, Sakr Y, Reinhart K, Investigators EIGo. International study of the prevalence and outcomes of infection in intensive care units. *JAMA* 2009; 302: 2323-2329.
38. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 2015; 161: 1202-1214.
39. Aneja R, Carcillo J. Differences between adult and pediatric septic shock. *Minerva Anestesiol* 2011; 77: 986-992.
40. Wheeler DS, Wong HR, Zingarelli B. Pediatric Sepsis - Part I: "Children are not small adults!". *Open Inflamm J* 2011; 4: 4-15.

41. Wong HR, Cvijanovich NZ, Anas N, Allen GL, Thomas NJ, Bigham MT, Weiss SL, Fitzgerald J, Checchia PA, Meyer K, Shanley TP, Quasney M, Hall M, Gedeit R, Freishtat RJ, Nowak J, Shekhar RS, Gertz S, Dawson E, Howard K, Harmon K, Beckman E, Frank E, Lindsell CJ. Developing a clinically feasible personalized medicine approach to pediatric septic shock. *Am J Respir Crit Care Med* 2015; 191: 309-315.
42. del Fresno C, Garcia-Rio F, Gomez-Pina V, Soares-Schanoski A, Fernandez-Ruiz I, Jurado T, Kajiji T, Shu C, Marin E, Gutierrez del Arroyo A, Prados C, Arnalich F, Fuentes-Prior P, Biswas SK, Lopez-Collazo E. Potent phagocytic activity with impaired antigen presentation identifying lipopolysaccharide-tolerant human monocytes: demonstration in isolated monocytes from cystic fibrosis patients. *J Immunol* 2009; 182: 6494-6507.
43. Pena OM, Pistolic J, Raj D, Fjell CD, Hancock RE. Endotoxin tolerance represents a distinctive state of alternative polarization (M2) in human mononuclear cells. *J Immunol* 2011; 186: 7243-7254.
44. Conner SD, Schmid SL. Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J Cell Biol* 2002; 156: 921-929.
45. Veiga E, Guttman JA, Bonazzi M, Boucrot E, Toledo-Arana A, Lin AE, Enninga J, Pizarro-Cerda J, Finlay BB, Kirchhausen T, Cossart P. Invasive and adherent bacterial pathogens co-Opt host clathrin for infection. *Cell Host Microbe* 2007; 2: 340-351.
46. Billingsley ML, Yun J, Reese BE, Davidson CE, Buck-Koehntop BA, Veglia G. Functional and structural properties of stannin: roles in cellular growth, selective toxicity, and mitochondrial responses to injury. *J Cell Biochem* 2006; 98: 243-250.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Figure legends

Fig. 1: Transcriptomic sepsis response signatures in FP. (A) Unsupervised hierarchical clustering analysis of 147 FP samples for the top 10% most variable probes (n=2716). **(B)** First three principal components plotted with the proportion of variance explained by each component, with samples colored by SRS group membership assigned by hierarchical and *k*-means clustering. **(C)** Volcano plot of differentially expressed probes for SRS1_FP vs SRS2_FP (red coloring shows fold change > 1.5, false discovery rate < 0.05). **(D)** Enriched functions, disease phenotypes, and predicted upstream regulators derived from differentially expressed probes in FP SRS groups and compared with the previously published CAP dataset. Enrichment was also seen for an endotoxin tolerance gene expression signature we previously defined(5) from publically available datasets(42, 43) in SRS1_FP relative to SRS2_FP tested using ROAST, a gene set enrichment test ($P<1x10^{-5}$) (**E**) Kaplan-Meier survival plot by SRS group (95% confidence intervals shaded) with a single sample selected at random for those patients with multiple samples to assign SRS membership. **(F)** The performance of the SRS group assignment models (gene sets) derived using and tested in the FP and previously published CAP datasets are shown by receiver operator characteristic (ROC) curves and the area under the curve (AUC) given for each. **(G)** Time course of patient SRS-FP group membership using serial samples and days from disease onset. **(H)** Expression of *CD163* over time from disease onset in samples from the 11 FP patients who move between SRS groups. Each point represents a sample, colored according to SRS group assignment, with lines linking samples from the same patient.

Fig. 2: Transcriptomic response to sepsis. (**A**) First two principal components of gene expression data plotted with the proportion of variance explained by each component shown.

Filled circles represent FP discovery samples (n=94), open circles represent CAP discovery samples (n=127). Samples are colored according to SRS group assignment, showing that sepsis response states show considerable overlap between sources of infection. Controls (n=10) are indicated by crosses. (**B**) Heatmap showing correlation between the first six principal components, SRS and clinical covariates for sepsis samples (FP n=94, CAP n=127). **(C)** Venn diagram showing the overlap in differential gene expression vs. controls in the sepsis response and the response to trauma (FDR < 0.05, FC > 1.5; first five days). Selected condition-specific enriched pathways and biological functions are noted. **(D)** Volcano plot of differentially expressed probes for FP (n=94) (left) and CAP (n=127) (right) vs controls (n=10) (red coloring shows FC > 1.5, FDR < 0.05).

Fig. 3: Variation in the sepsis transcriptome according to source of infection. (**A**) Venn diagram showing the overlap in the FP and CAP sepsis response in terms of the number of differentially expressed genes vs. controls (FP n=64 samples, CAP n=73 samples, Controls n=10 samples). (**B**) Correlation of differential gene expression between sepsis due to FP (n=64 samples) and controls (n=10 samples), and sepsis due to CAP (n=73 samples) and controls (n=10 samples). (**C**) Volcano plot of differentially expressed probes for CAP vs FP (red coloring shows FC >1.5, FDR <0.05, positive fold change indicates relative upregulation in CAP). (**D**) Enriched pathways and predicted upstream regulators derived from differentially expressed genes for CAP vs FP in the discovery (CAP n=73 samples, FP n=64 samples) and validation cohorts (CAP n=53 samples, FP n=53 samples). (**E**) Most significantly enriched network (P $1x10^{-39}$) on comparison of CAP vs FP with differentially expressed genes shaded (red upregulation, green downregulation) and log fold change with P value shown for each gene. (**F**) Volcano plot of the

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

differentially expressed probes for positive viral diagnosis (n=25) vs no positive viral diagnosis (n=240) in 265 CAP patients. Red coloring indicates FC >1.5, FDR <0.05.

Fig. 4: Dynamics of gene expression in sepsis due to FP and CAP. (**A**) Volcano plot of the differentially expressed probes for Day 1 vs 5 samples from the same patient in FP (n=9) (left) and CAP (right) (n=17) (red coloring indicates FC >1.5 and FDR <0.05). (**B)** Examples of genes showing significant temporal variation in expression over time for patients with sepsis due to FP included the kinase activator *CDK5R1* and complement component *C1QB*; other examples are the adhesion receptor *EMR3* and cysteine protease *CAPN13*. (**C**) Examples of genes showing variation in expression over time for patients with sepsis due to CAP included cell cycle genes *NIT2* and the hematopoiesis regulatory transcription factor gene *MYB*; other examples are the antimicrobial neutrophil peptidase CTSG, inflammatory regulator *CAMP*, defensin *DEFA4* and neutrophil granule elastase *ELANE* (**D**) Further examples of genes differentially expressed in FP over time plotted from time of onset of FP include the AP2 associated kinase 1 *AAK1* involved in regulating clathrin-mediated endocytosis (44), critical for bacterial entry (45); and *SNN* encoding stannin, a mitochondrial damage sensor (46).

	Discovery cohort			Validation cohort		
	CAP (n=73)	**FP (n=64)**	**p value**	**CAP (n=53)**	**FP (n=53)**	**p value**
Age (years)	60.6 (15.6)	65.1 (17.5)	NS	68.4 (13.7)	68 (12.7)	NS
Male sex	37 (50.6%)	28 (43.7%)	NS	37 (69.8%)	31 (58.5%)	NS
APACHE II score	18 (6.2)	15 (5.9)	0.005	19.7 (6.1)	15.7 (5.9)	0.0008
SOFA score	6.8 (3.5)	5.4 (4.0)	0.03	6.2 (3.8)	6.3 (3.5)	NS
Mortality						
14 days	8 (11.0%)	6 (9.4%)	NS†	7 (13.2%)	6 (11.3%)	NS†
28 days	13 (17.8)	8 (12.5%)	NS†	7 (13.2%)	7 (13.2%)	NS†
6 months	17 (23.3%)	11 (17.2%)	NS†	11 (20.8%)	12 (22.6%)	NS†
Infection						
Gram-positive bacteria	14 (19.2%)			7 (13.2%)		
Gram-negative bacteria	8 (11.0%)			5 (9.4%)		
Viral	6 (8.2%)			5 (9.4%)		
Mechanical Ventilation	51 (69.9%)	34 (53.1%)	0.05	19 (35.8%)	21 (39.6%)	NS
Respiratory Rate	29.2 (9.2)	20.0 (8.2)	<0.0001	31.1 (8.7)	22.7 (9.0)	<0.0001
Days of Respiratory Support	9.6 (13.5)	4.7 (8.1)	0.01	8.8 (12.1)	9.3 (14.8)	NS
Oxygenation Index	19.3 (9.9)	30.9 (12.3)	<0.0001	21.2 (7.5)	28.2 (12.2)	<0.0001
Vasopressors/inotropes			NS			NS
No dose	37 (50.7%)	35 (54.7%)		36 (67.9%)	27 (50.9%)	
Low dose	1 (1.4%)	1 (1.6%)		0 (0%)	0 (0%)	
Medium dose	10 (13.7%)	5 (7.8%)		7 (13.2%)	6 (11.3%)	
High dose	24 (32.9%)	23 (35.9%)		10 (18.9%)	20 (37.7%)	
Mean arterial pressure (lowest, mmHg)	65.0 (12.8)	67.7 (12.5)	NS	69.6 (11.1)	67.5 (14.6)	NS
Temperature (low)	36.4 (0.8)	36.1 (0.7)	0.018	36.0 (1.0)	36.1 (0.6)	NS
Hematocrit	35.4 (6.9)	30 (5.9)	<0.0001	35.0 (7.9)	31.5 (6.6)	0.02
Proportion of leukocytes						
Lymphocytes	9.17%	6.52%	0.008	9.05%	8.77%	NS
Polynucleocytes	84.10%	87.98%	0.029	83.40%	85.30%	NS
Mononucleocytes	6.73%	5.50%	NS	7.51%	5.95%	NS

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Bicarbonate	25.0 (6.8)	22.1 (5.1)	0.008	24.6 (6.8)	23.1 (5.0)	NS
Renal replacement therapy	6 (8.2%)	5 (7.8%)	NS	5 (9.4%)	8 (15.1%)	NS

Table 1: Comparison of clinical characteristics of sepsis patients based on source of infection in the discovery and validation cohorts. Data are n (%) or mean (SD) unless otherwise specified. APACHE II = Acute Physiology and Chronic Healthy Evaluation II. SOFA=Sequential Organ Failure Assessment on day of sampling. Statistical analysis t test unless otherwise specified. $^{*}\chi^2$ test. †Log-rank test.

Number of patients	117
Age (mean years ± SD)	66.4 (±15.5)
Sex, male (%)	50.4%
APACHE II score	15 (2-40)
SOFA score	5 (0-18)
Mortality	
14 day mortality	12 (10.3%)
28 day mortality	15 (12.8%)
6 month mortality	23 (19.7%)
Cause of peritonitis	
Diverticular disease	30 (25.6%)
Surgical anastomosis breakdown	28 (23.9%)
Malignancy	14 (12%)
Trauma	3 (2.6%)
Other/unknown	42 (35.9%)
Mechanical ventilation	55 (47%)
Vasopressors	55 (47%)
Renal replacement therapy	13 (11%)
Mean arterial pressure (mmHg)	68.2 (14.8)

Table 2. Characteristics of fecal peritonitis patients (all patients) on day 1.

Data are n (%) unless otherwise specified. APACHE II=Acute Physiology and Chronic Health Evaluation II, median (range). SOFA= Sequential Organ Failure Assessment on day of sampling, median (range). Arterial pressure, lowest mmHg (SD).

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

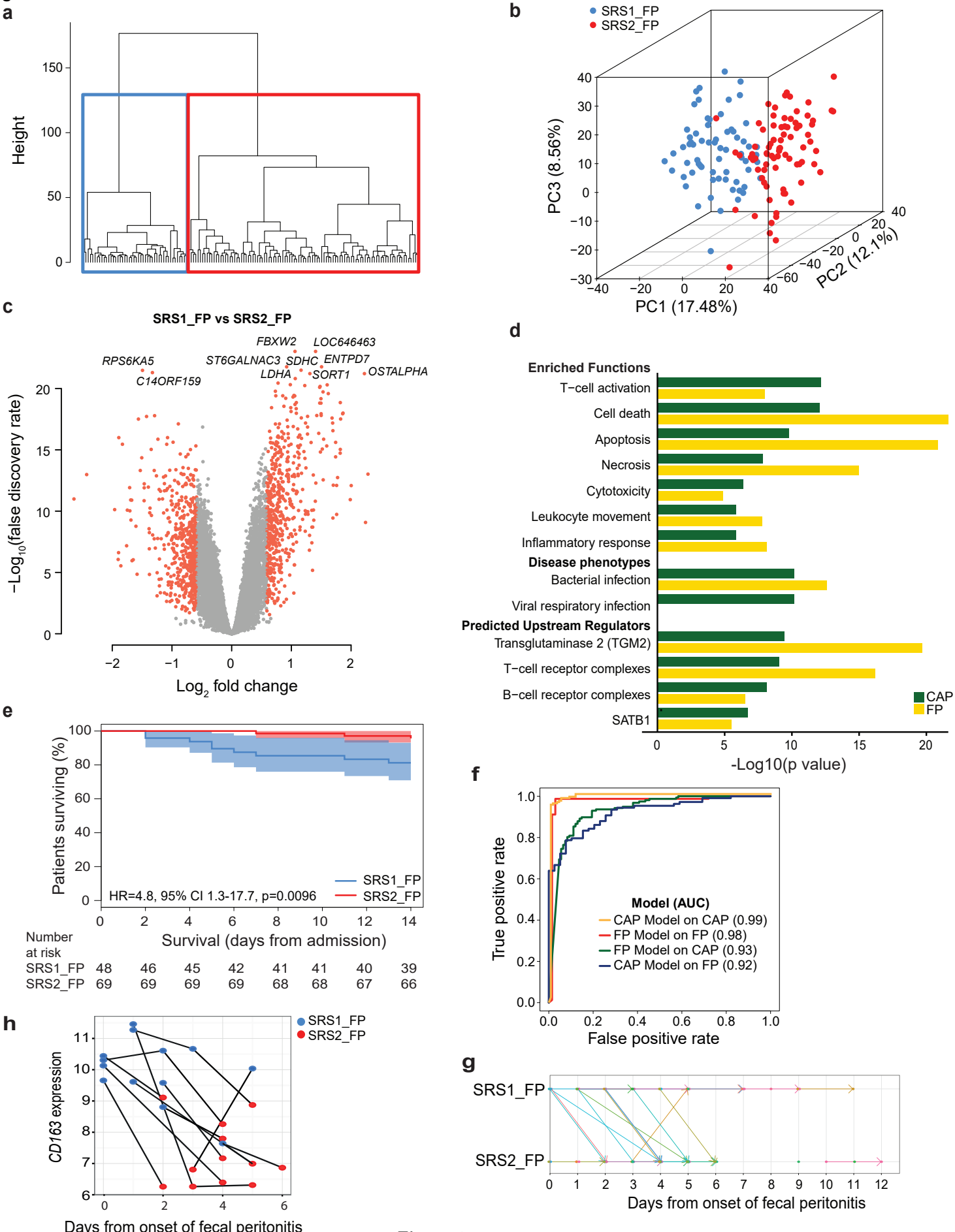


Figure 1

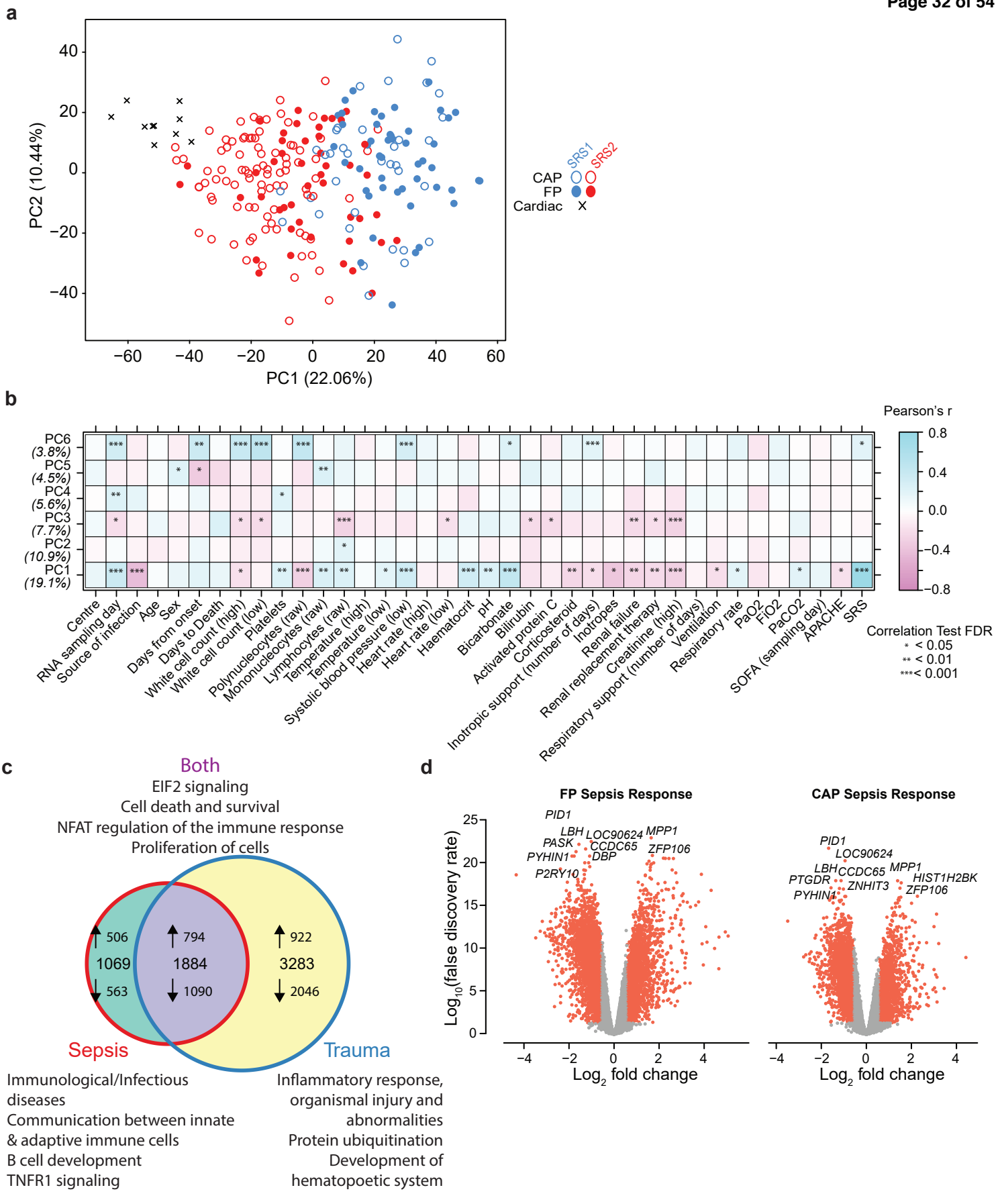


Figure 2

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

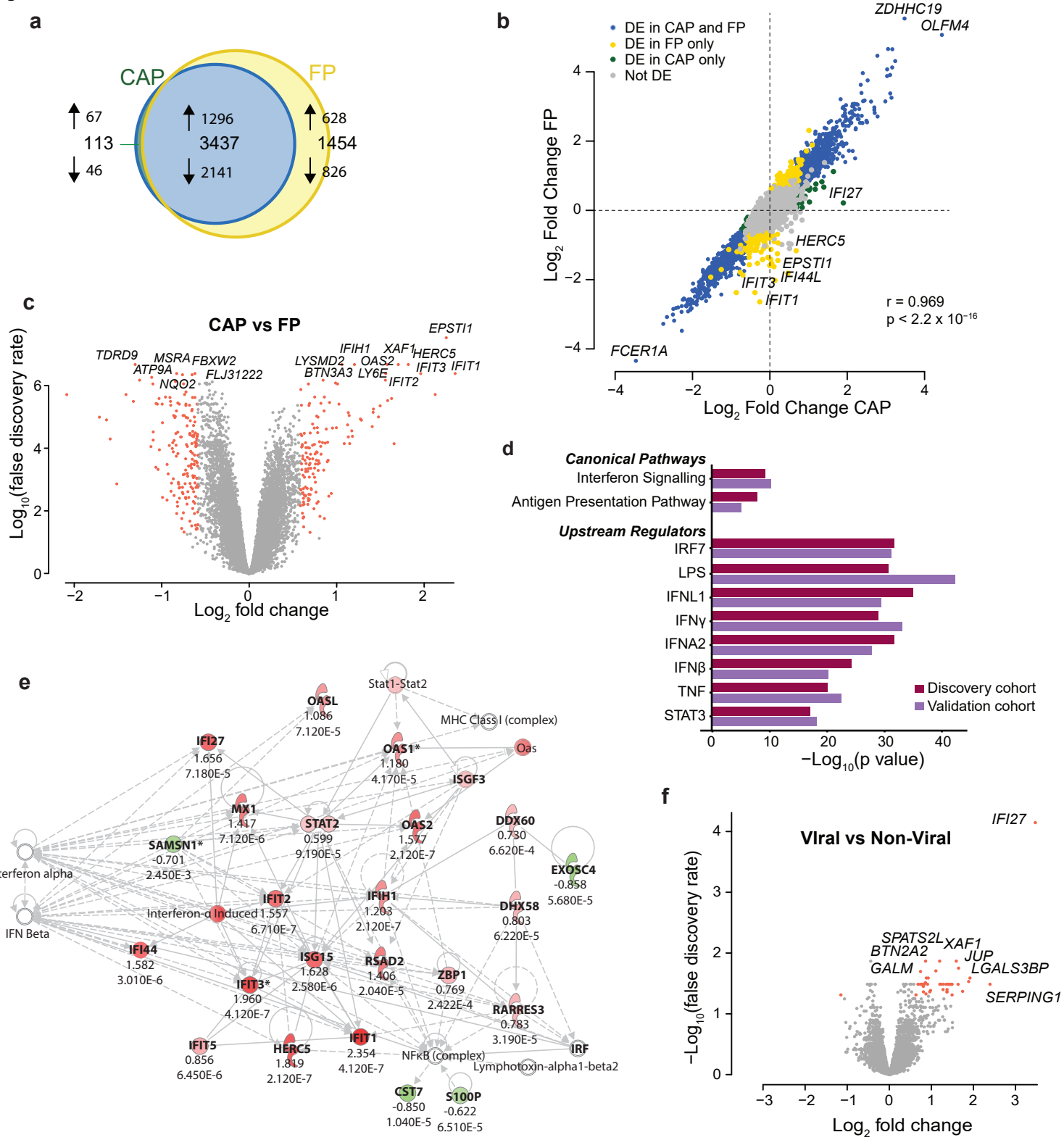


Figure 3

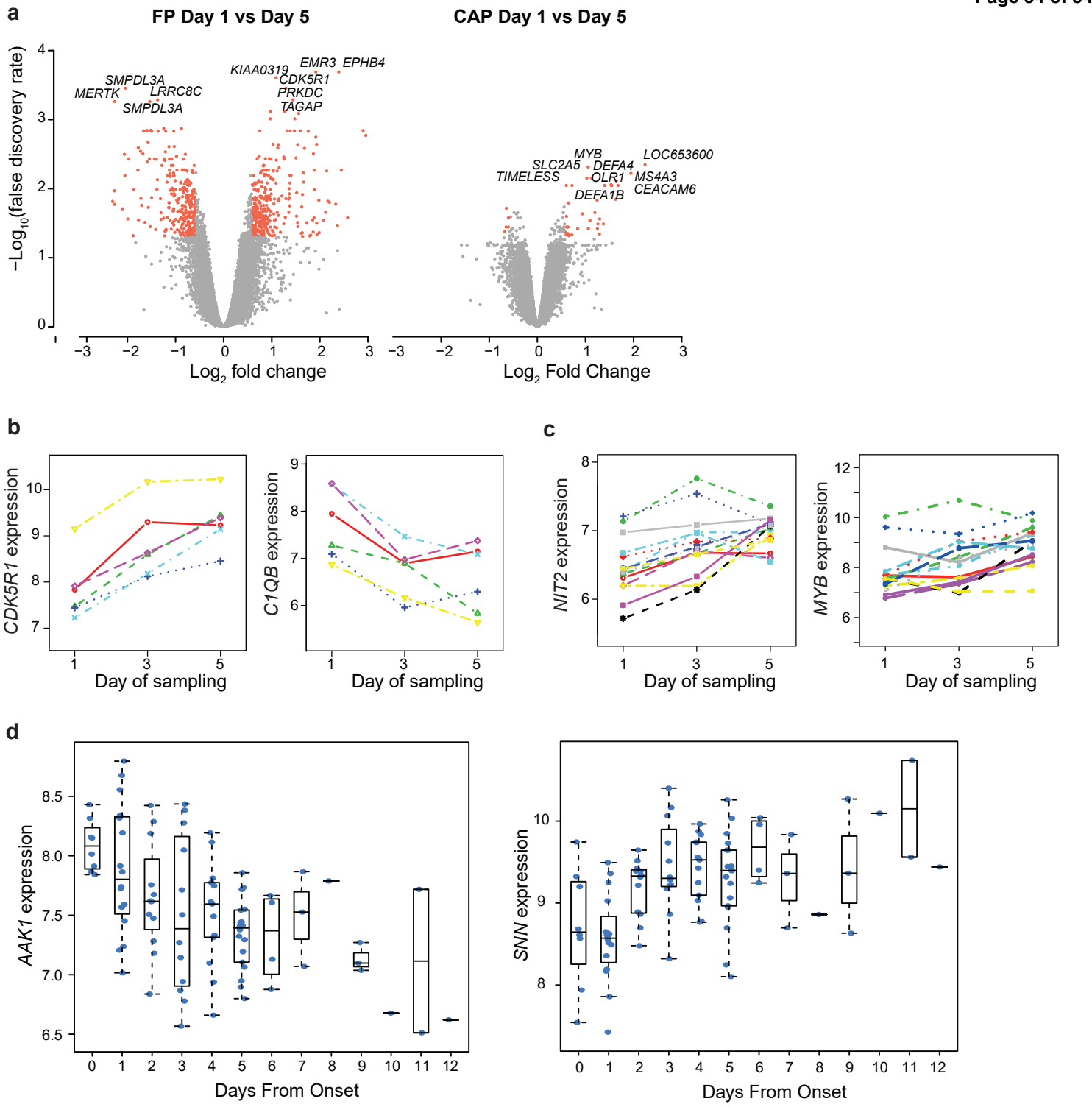


Figure 4

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Online Data Supplement

Shared and distinct features of the sepsis transcriptomic response to fecal peritonitis and pneumonia

Authors: Katie L. Burnham, Emma E. Davenport, Jayachandran Radhakrishnan, Peter Humburg, Anthony C. Gordon, Paula Hutton, Eduardo Svoren-Jabalera, Christopher Garrard, Adrian V.S. Hill, Charles J. Hinds, Julian C. Knight.

Supplementary Methods

Supplementary References

Supplementary Figures

Figure E1. Schematic of the composition of the discovery and validation cohorts.

Figure E2. Correlation of differential gene expression between SRS groups in FP and CAP.

Figure E3. Principal component analysis of FP samples from the discovery and validation cohorts showing SRS group assignment

Figure E4. Comparison of the SRS groups to pediatric sepsis endotypes.

Figure E5. Expression of SRS_FP model genes in patients moving between SRS groups.

Figure E6. Correlation between global gene expression and clinical covariates.

Figure E7. Histogram of results of CAP vs FP differential expression analysis following permutation.

Figure E8. Effect of cell count on differential gene expression analysis between CAP and FP.

Figure E9. ROC curves for gene set discriminating between CAP and FP.

Figure E10. Histogram of RNA sampling day for CAP and FP samples.

Figure E11. Principal component analysis of FP samples from the discovery and validation cohorts combined using ComBat.

Supplementary Tables

Table E1. Supplementary data (Excel file)

Table E2. Participating hospitals involved in patient recruitment and GAinS investigators.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Supplementary Methods

Patient recruitment and exclusion criteria

Adult patients (>18y) were recruited at 25 UK ICUs between 2006 and 2013 as part of the UK Genomic Advances in Sepsis (GAinS) study (www.ukccg-gains.org) (1, 2). Ethics approval was granted nationally (REC Reference Number 05/MRE00/38 and 08/H0505/78) and for individual participating centers, with informed consent obtained from patients or their legal representative. Sepsis was diagnosed according to ACCP/SCCM guidelines, and all patients showed evidence of organ dysfunction (1). FP was diagnosed at laparotomy as inflammation of the peritoneal membrane secondary to large bowel perforation and fecal contamination (3). CAP was defined as a febrile illness associated with a cough, sputum production, breathlessness, leukocytosis and radiological features of pneumonia, acquired in the community or within two days of ICU admission (4, 5). Exclusion criteria: patient or legal representative unwilling or unable to give consent; age <18 years; pregnancy; advanced directive to withhold or withdraw life sustaining treatment; admission for palliative care only; or immune-compromise. Demographics and clinical covariates were recorded using an electronic case report form which included details of the results of microbiological investigations as previously described (1). Equivalent samples obtained from 10 patients scheduled to undergo elective cardiac surgery (coronary artery bypass grafting/valve replacement/ cardiopulmonary bypass) were used to represent baseline gene expression in non-septic individuals. The sample size for the validation cohort was calculated based on the original expression data with a group size of 53 giving >90% power to detect a 1.5-fold change in gene expression at a false discovery rate of 0.05. The primary endpoint for the SRS groups was prospectively selected as mortality at 14 days from ICU admission based on findings in CAP patients (1).

RNA extraction

RNA was extracted from the leukocyte samples (stored at -80^{o}C) using the Total RNA Isolation Protocol (Ambion). The contents of the filters were collected following lysis with a guanidine thiocyanate-based solution. Cellular proteins and DNA were degraded using Proteinase K and DNase I respectively and RNA purified using magnetic bead technology. The RNA thus obtained was quantified by spectrophotometry (NanoDrop) and quality checks were carried out on a small subset of the samples by on-chip electrophoresis (BioRad Bioanalyzer).

Microarray analysis

Sepsis samples were processed as two cohorts of 227 and 108 samples, with ten samples from patients scheduled to undergo elective cardiac surgery also included in the first batch. The discovery cohort consists of all samples available for each patient recruited pre-August 2010 that passed quality control checks. For the validation cohort, samples from the first available time point with high quality RNA were selected for 54 FP patients and 54 approximately age and sex matched CAP patients recruited subsequent to the first microarray analysis. The FP samples from the discovery and validation cohorts were combined using ComBat (6) (Supplementary Fig. E11). Gene expression data for validation cohort CAP samples and the majority of discovery cohort CAP samples were published in Davenport *et al*, 2016 (1). Discovery cohort samples were re-run in this study to allow for multiple time points per individual. Microarrays for FP and CAP validation cohorts were run at the same time. The traumatic injury and pediatric datasets were accessed through NCBI GEO database (accession: GSE36809, GSE26440). The pediatric

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

expression data was log transformed and in parallel re-normalized using the robust multi-array average (RMA) for comparative differential expression analysis.

Raw data were prepared using Illumina's Genomestudio. Data backgrounds were subtracted and probes with a detection p-value <0.05 in at least 5% samples retained for downstream analysis in the discovery cohort (27,159 probes). The raw data were transformed and normalized using the vsn package in R (7). The validation cohort was normalized against the discovery cohort, with other QC steps conducted independently. Following QC including principal component analysis, 6 samples were removed from the discovery cohort and two from the validation, leaving 221 and 106 samples respectively.

Statistical analysis

Demographic and phenotypic data were compared with t-tests for continuous data, and Chi-square tests for count data. Power calculations were performed using the R package sizepower. Differential gene expression was assessed using the limma package (8), which fits a generalized linear model to the expression of each gene and employs an empirical Bayes approach to take into account the overall variance in the data set. Genes with an FDR <0.05 and a FC >1.5 were considered to be differentially expressed. Further investigation of longitudinal gene expression profiles was carried out with the timecourse package (9). The enrichment of a given gene set in a particular gene expression contrast (e.g. SRS1 vs SRS2 in FP and CAP) was tested using ROAST, a rotation-based gene set enrichment test that incorporates effect size and directionality (10). Biological pathway enrichment analysis was carried out with Ingenuity Pathway Analysis (IPA - QIAGEN, www.qiagen.com/ingenuity QIAGEN Redwood City). Enrichment or over-

representation of different processes is based on the number of genes identified and the total number of genes known to be associated with that process in the reference set (IPA Knowledge Base, experimentally observed human molecules and/or relationships) with p value calculated using a right-tailed Fisher exact test. Predictors for the SRS_FP and diagnostic models were selected using GeneRave (11) from differentially expressed genes with moderate to high expression (expression > 6.5) in at least the number of samples in the smaller group. GeneRave uses a sparse regression method to select a minimal number of explanatory variables from a large number of potential predictors. Patient subgroups were resolved using the FactoMineR package, which carries out agglomerative hierarchical clustering using Ward's method with Euclidean distance used to initiate the appropriate number of groups, followed by *k*-means clustering to consolidate group membership. Plotting the within group sum of squares was used to determine an appropriate number of clusters. Comparisons to the SRS groups previously observed in CAP patients were carried out using a seven gene signature (1). Survival differences were assessed by log rank test and cox proportional hazards regression models using the R package survival.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

References

1. Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, Rautanen A, Gordon AC, Garrard C, Hill AVS, Hinds CJ, Knight JC. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *The Lancet Respiratory medicine* 2016; 4: 259-271.
2. Rautanen A, Mills TC, Gordon AC, Hutton P, Steffens M, Nuamah R, Chiche JD, Parks T, Chapman SJ, Davenport EE, Elliott KS, Bion J, Lichtner P, Meitinger T, Wienker TF, Caulfield MJ, Mein C, Bloos F, Bobek I, Cotogni P, Sramek V, Sarapuu S, Kobilay M, Ranieri VM, Rello J, Sirgo G, Weiss YG, Russwurm S, Schneider EM, Reinhart K, Holloway PA, Knight JC, Garrard CS, Russell JA, Walley KR, Stuber F, Hill AV, Hinds CJ, Investigators EEG. Genome-wide association study of survival from sepsis due to pneumonia: an observational cohort study. *The Lancet Respiratory medicine* 2015; 3: 53-60.
3. Tridente A, Clarke GM, Walden A, McKechnie S, Hutton P, Mills GH, Gordon AC, Holloway PA, Chiche JD, Bion J, Stuber F, Garrard C, Hinds CJ, Gen OI. Patients with faecal peritonitis admitted to European intensive care units: an epidemiological survey of the GenOSept cohort. *Intensive Care Med* 2014; 40: 202-210.
4. Angus DC, Marrie TJ, Obrosky DS, Clermont G, Dremsizov TT, Coley C, Fine MJ, Singer DE, Kapoor WN. Severe community-acquired pneumonia: use of intensive care services and evaluation of American and British Thoracic Society Diagnostic criteria. *Am J Respir Crit Care Med* 2002; 166: 717-723.
5. Walden AP, Clarke GM, McKechnie S, Hutton P, Gordon AC, Rello J, Chiche JD, Stueber F, Garrard CS, Hinds CJ. Patients with community acquired pneumonia admitted to European Intensive Care Units: an epidemiological survey of the GenOSept cohort. *Crit Care* 2014; 18: R58.
6. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; 8: 118-127.
7. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 2002; 18 Suppl 1: S96-104.
8. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; 3: Article3.
9. Tai YC, Speed TP. On gene ranking using replicated microarray time course data. *Biometrics* 2009; 65: 40-51.
10. Wu D, Lim E, Vaillant F, Asselin-Labat ML, Visvader JE, Smyth GK. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* 2010; 26: 2176-2182.
11. Kiiveri HT. A general approach to simultaneous model fitting and variable elimination in response models for biological data with many more variables than observations. *BMC Bioinformatics* 2008; 9: 195.

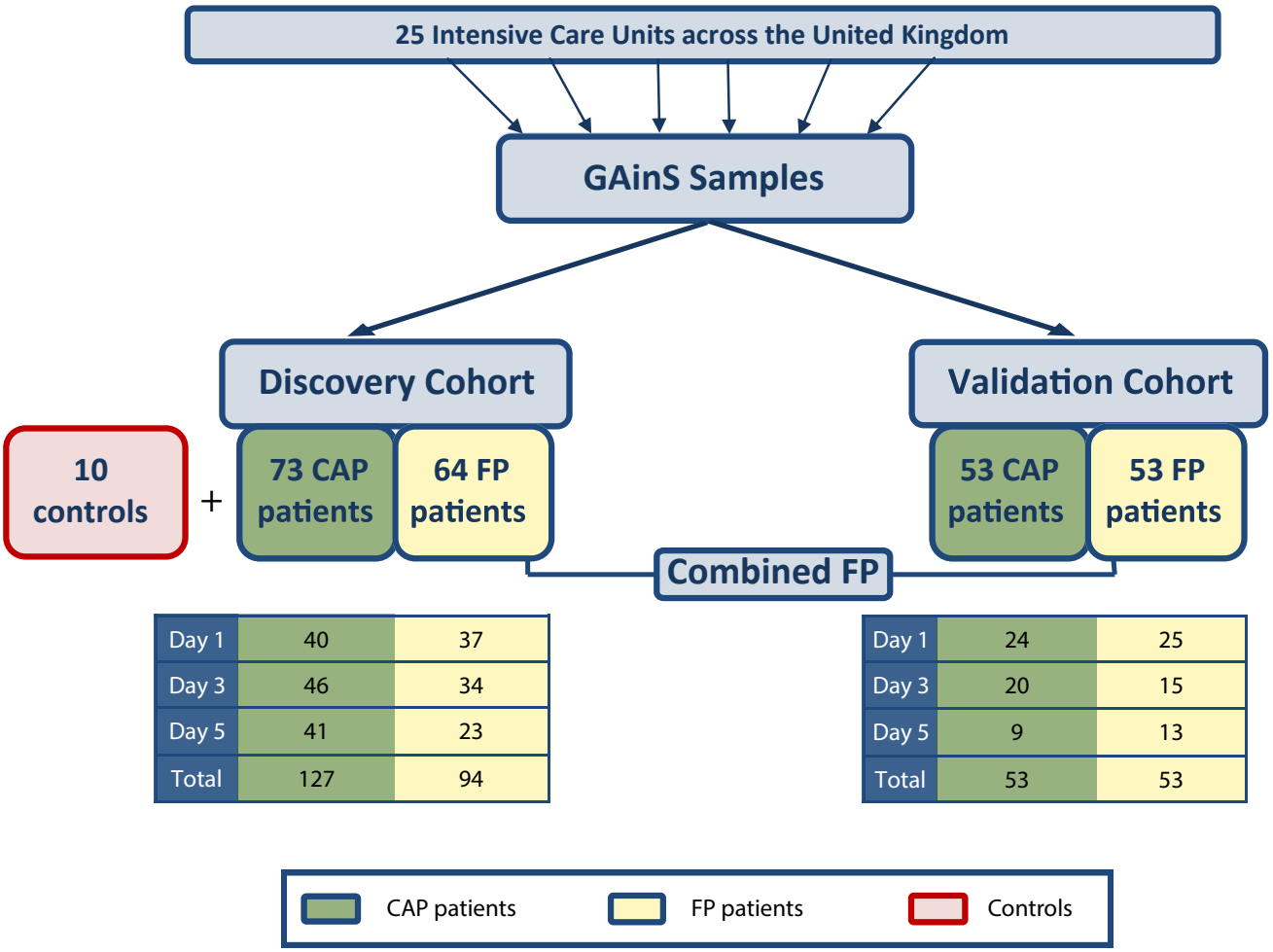


Figure E1. Schematic of the composition of the discovery and validation cohorts. Number of patients and samples included in the discovery and validation cohorts following quality control of gene expression data, categorised by aetiology. Controls are pre-operative patients scheduled to undergo elective cardiac surgery. For the investigation of sepsis response signatures in FP, all FP patients were used (117 patients).

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

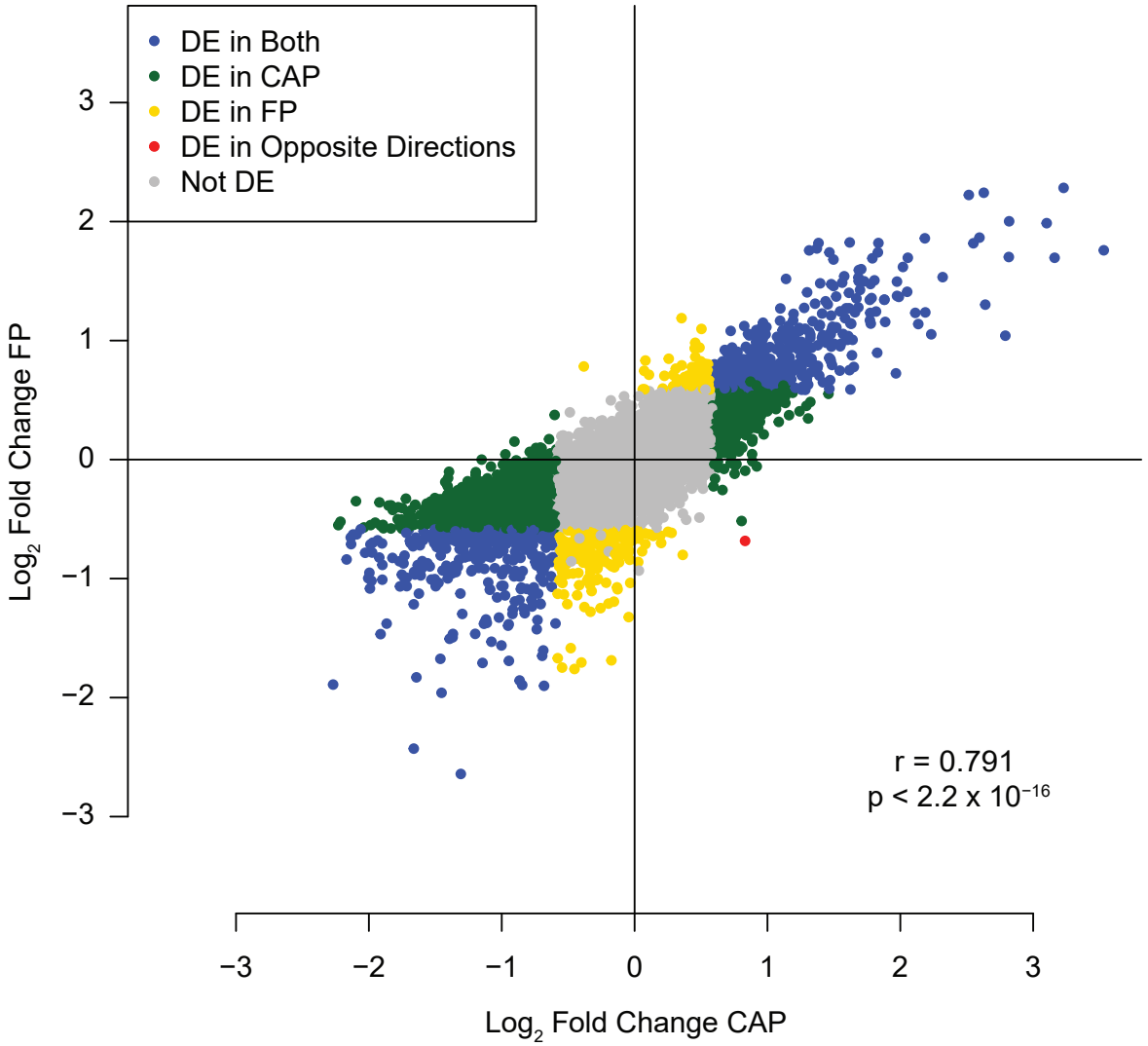


Figure E2. Correlation of differential gene expression between SRS groups in FP and CAP. The log fold change for differential expression of each measured gene between SRS1 and SRS2 in FP is plotted against log fold change for the same gene between SRS1 and SRS2 in CAP. The log fold changes are highly correlated (r 0.79, p <2.2x10^{-16}). Blue colour indicates significant differential expression between SRS groups in both CAP and FP;dark green in CAP only; yellow in FP only; red indicates differential expression in both but in opposite directions; and grey indicates that the gene is not significantly differentially expressed.

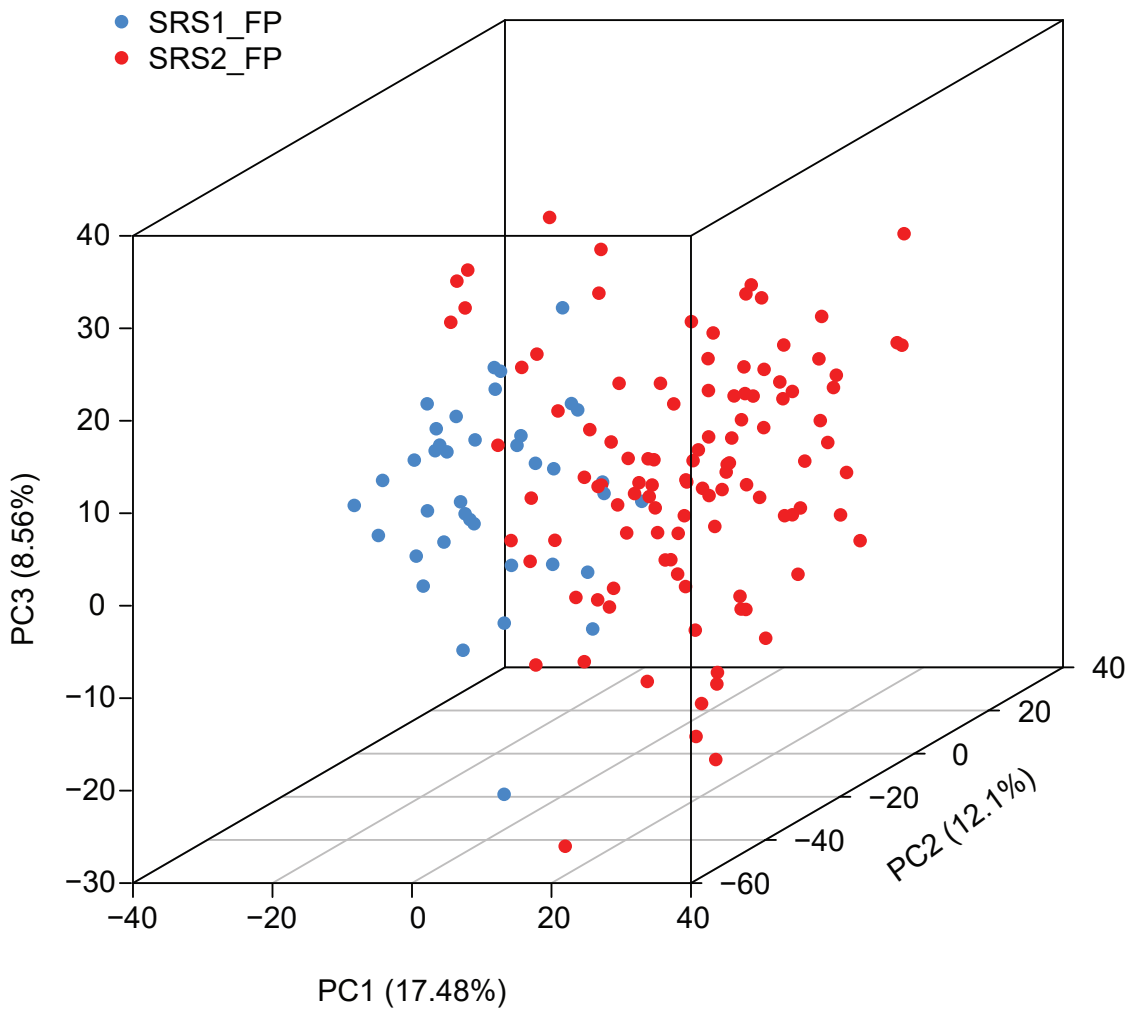


Figure E3. Principal component analysis of FP samples from the discovery and validation cohorts showing SRS group assignment.
First three principal components plotted with the proportion of variance explained by each component, with samples coloured according to SRS group membership assigned using a seven gene signature derived from the previously published CAP cohort.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

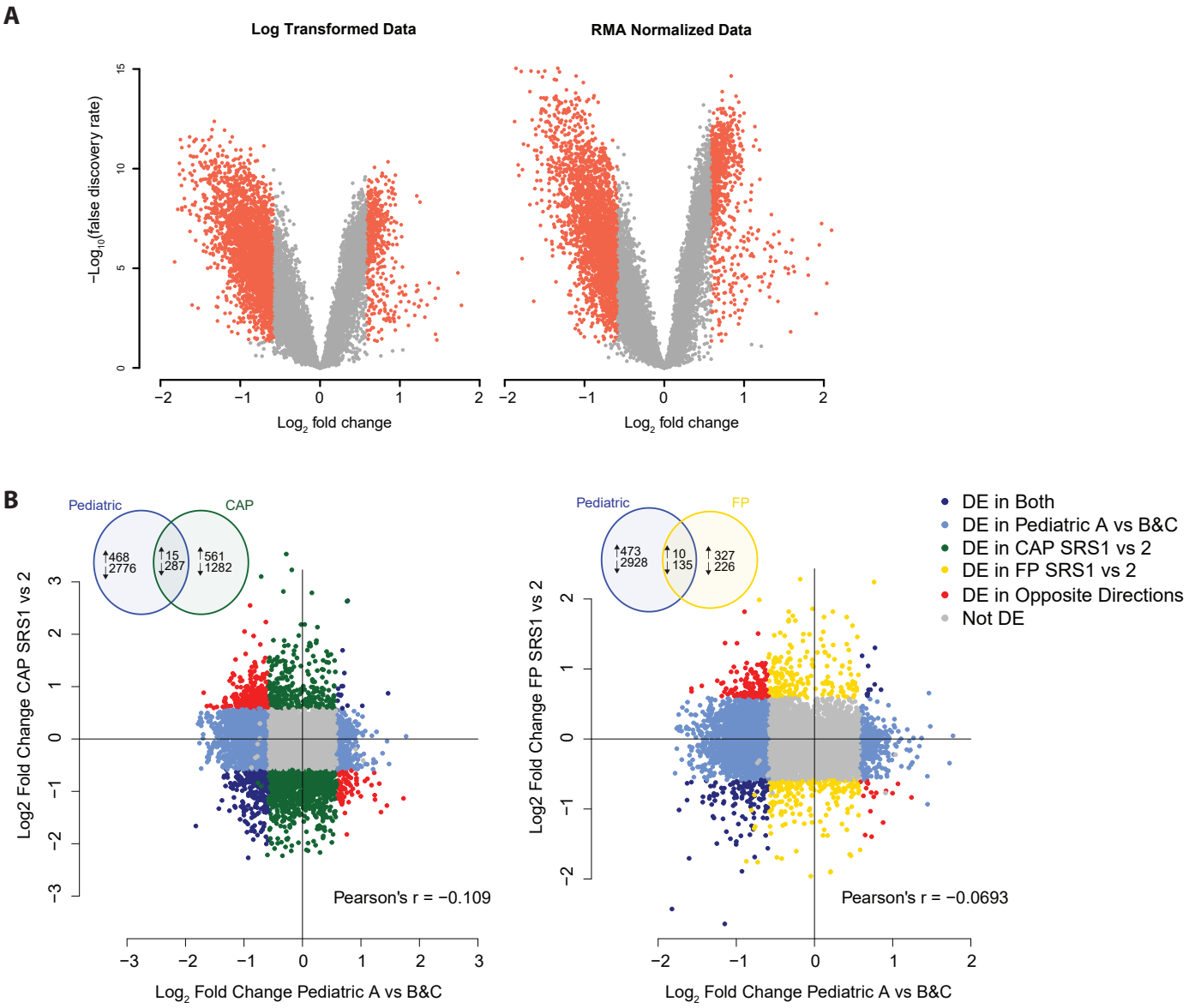


Figure E4. Comparison of the SRS groups to pediatric sepsis endotypes.
(A) Volcano plots of differential gene expression between pediatric endotype A (n=28), and endotypes B and C (n=70). The left hand plot shows the results for log transformed pre-processed data (GSE26440), and the right hand plot for RMA normalized data (raw data downloaded from GEO). Red coloring indicates probes with FDR < 0.05 and FC > 1.5. **(B)** Correlation plots of the differential gene expression outputs. On the left the fold changes of each gene for CAP SRS1 vs 2 is plotted against the fold changes for the same genes for the pediatric endotype comparison (using the log transformed data). On the right the results for the FP SRS contrast are plotted against the pediatric output. Points are colored according to whether the gene was significantly differentially expressed in one, both, or neither contrast. Above the plots are Venn diagrams showing the overlap in the genes differentially expressed between pediatric endotypes, and between the SRS groups. The number of genes up- and downregulated are given for each comparison, and only genes with expression data in all three cohorts are considered. The SRS genesets were found not to be significantly enriched in the pediatric contrast by ROAST gene set enrichment tests.

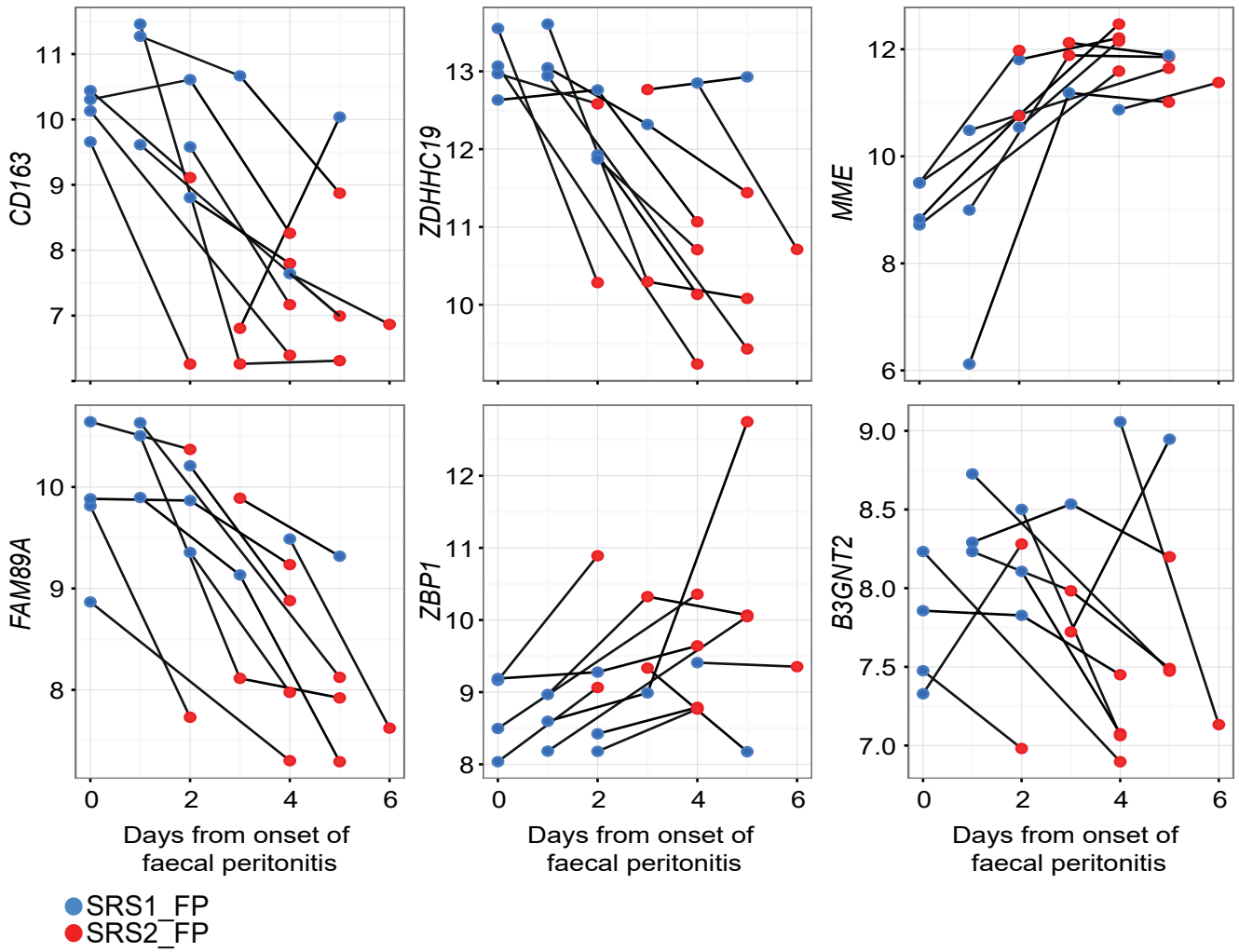


Figure E5. Expression of SRS_FP model genes in patients moving between SRS groups. Expression of the six genes selected as predictors of SRS_FP group membership (*CD163, ZDHHC19, MME, FAM89A, ZBP1,* and *B3GNT2*) in the 11 patients who move between SRS groups over time. Each point represents a sample, coloured by SRS group assignment and plotted according to time since disease onset. Samples from the same patient are linked, showing the changes in gene expression over time as patients move from one SRS group to another

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

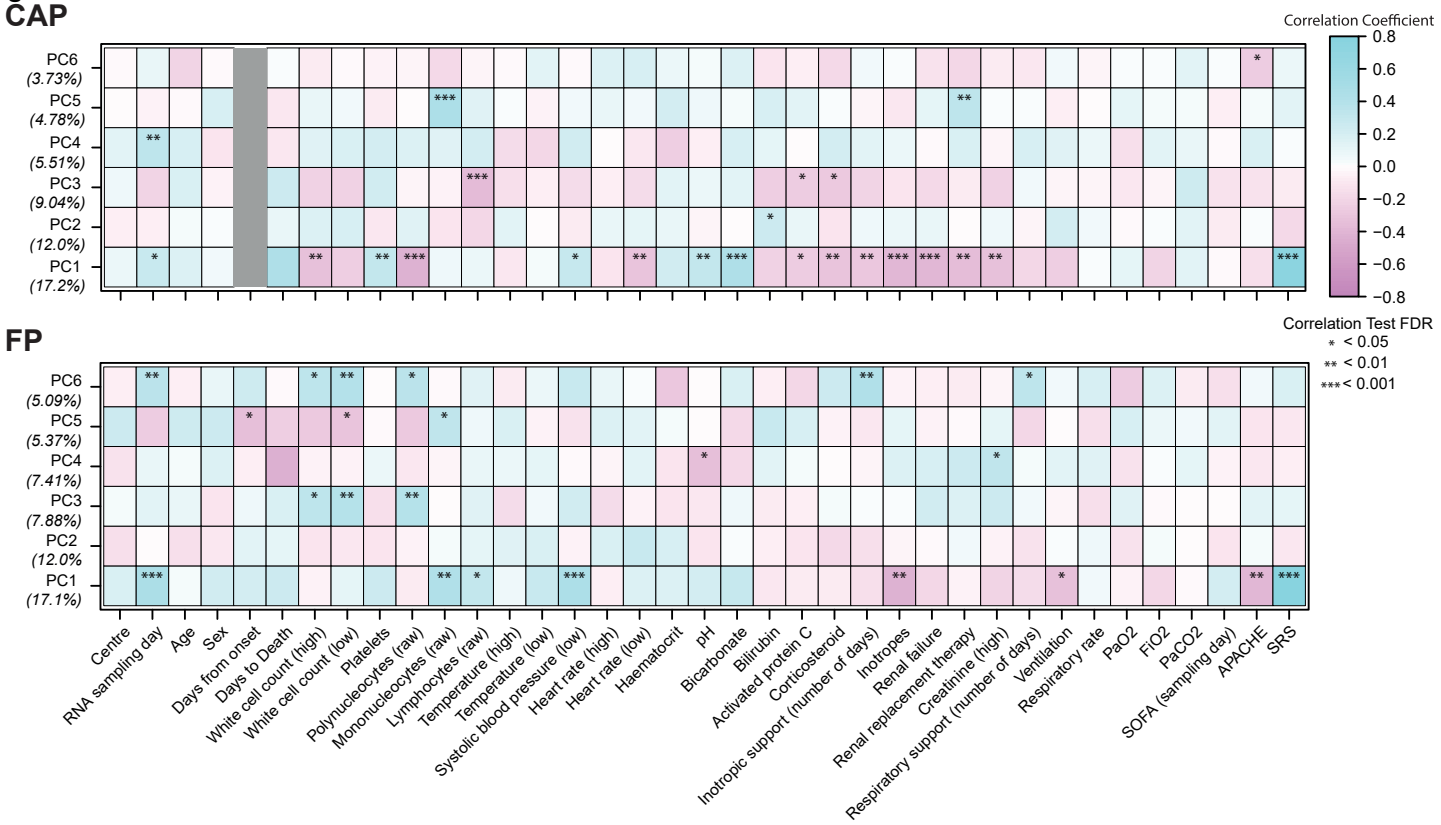
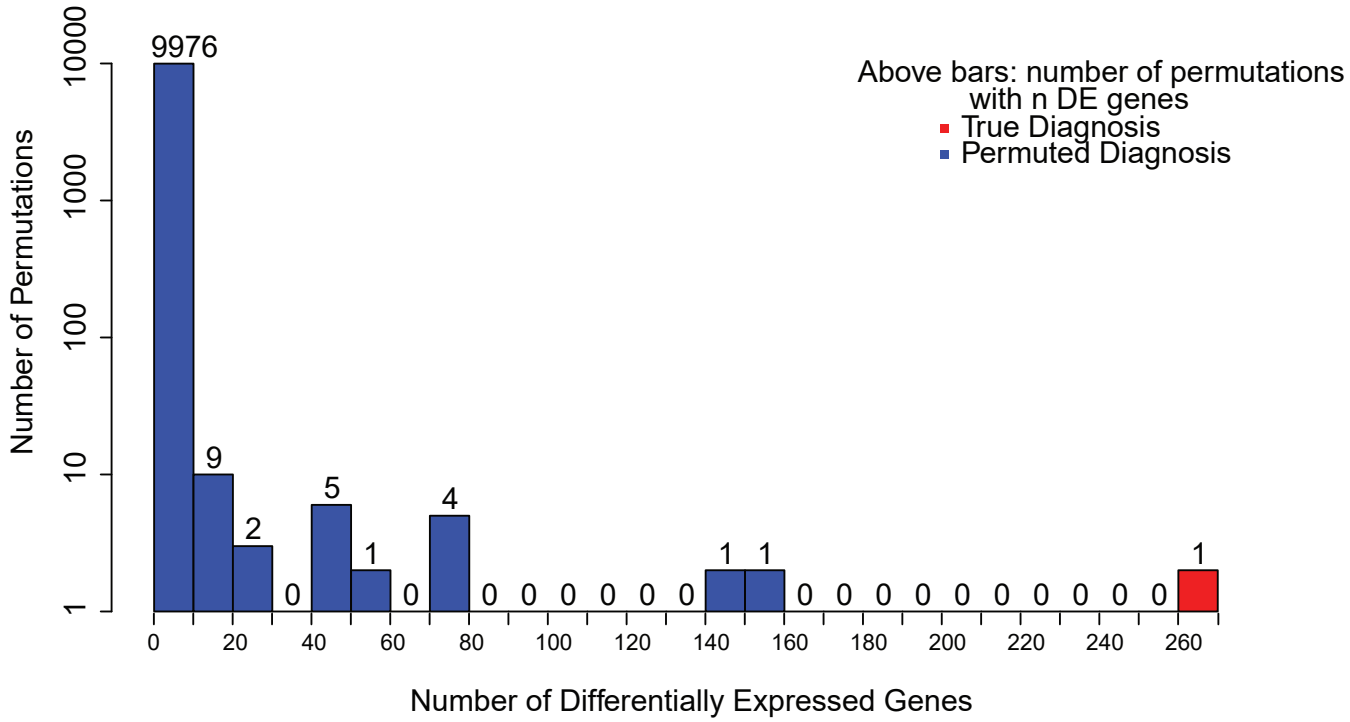


Figure E6. Correlation between global gene expression and clinical covariates.
Heatmaps showing correlation between clinical covariates and the first six principal components of gene expression for CAP (top) and FP (bottom) samples.



Supplementary Figure E7. Histogram of results of CAP vs FP differential expression analysis following permutation.
In order to determine if the number of genes differentially expressed between CAP and FP patients was more than expected by chance, the diagnostic labels were permuted and the differential expression analysis repeated 10,000 times. In each case, the number of differentially expressed genes (FDR < 0.05, FC >=1.5) was recorded. The comparison for the true diagnostic labels had significantly more differentially expressed genes (n=263) than would be expected by chance (p=0.0001). The number of permutations returning a given number of differentially expressed genes is plotted as a histogram, with the number of permutations for each block also shown above each bar. The result for the true instance is colored in red.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

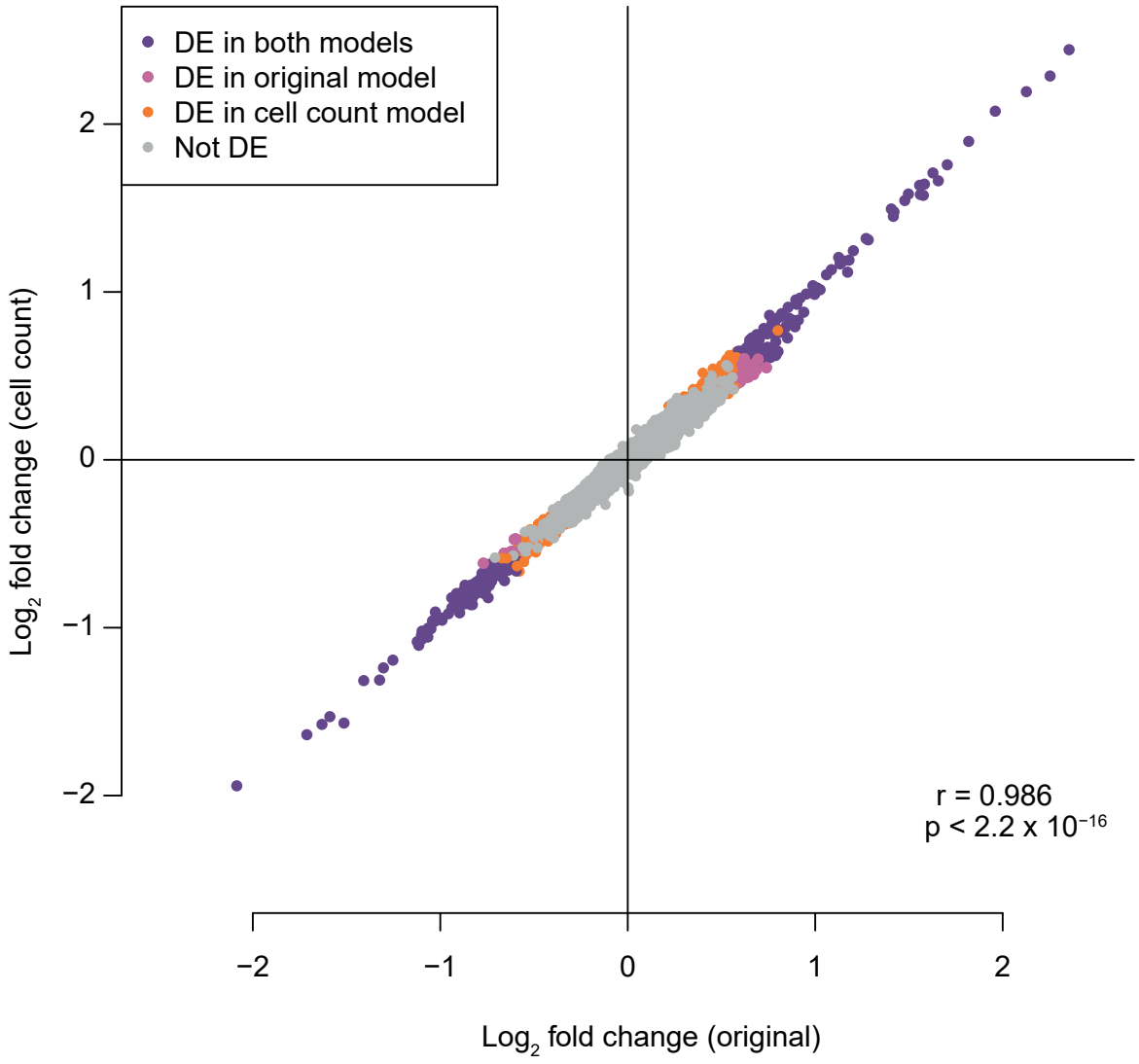


Figure E8. Effect of cell count on differential gene expression analysis between CAP and FP. The log fold change for differential gene expression of all measured genes between CAP and FP accounting for mononucleocyte, polynucleocyte, and lymphocyte count is plotted against the log fold change for the same genes between CAP and FP in the minimal model. The log fold changes are highly correlated (r 0.986, $p < 2.2 \times 10^{-16}$). Purple colouring indicates significant differential expression in both models; pink in the original model only; orange in the model including cell count only; and grey indicates the gene is not significantly differentially expressed between CAP and FP in either model.

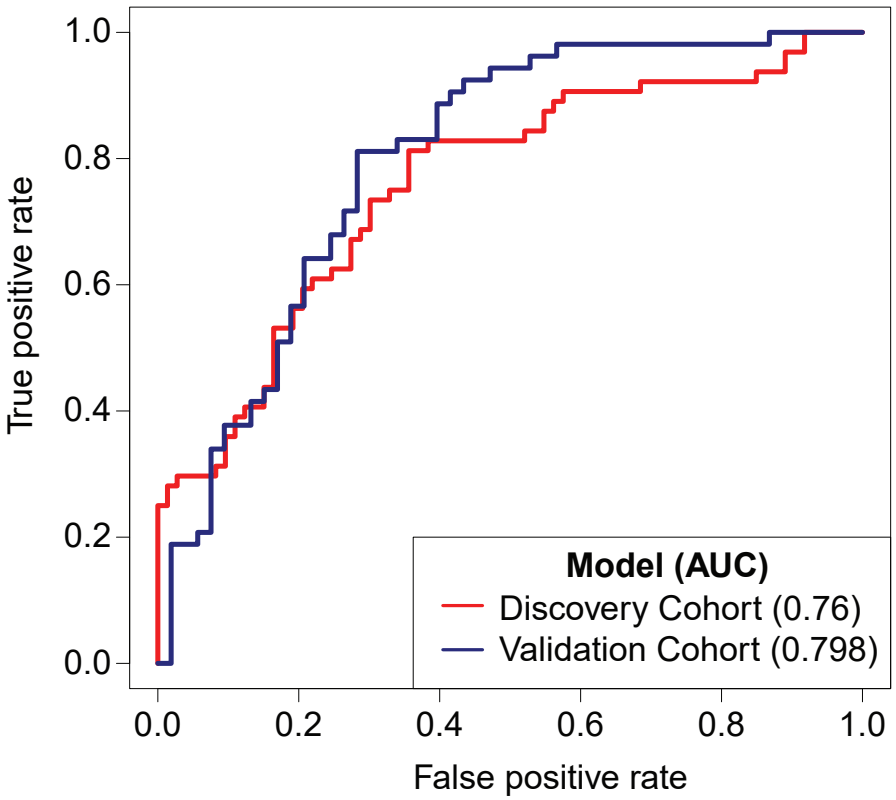


Figure E9. ROC curves for gene set discriminating between CAP and FP.
A subset of genes discriminating between CAP and FP samples were identified, comprising *EPHB1*, *NQO2*, *ARG1*, *HMGB2*, *FGL2* and *GPR162.* The performance of the model was tested by leave-one-out validation in the discovery cohort and on all samples in the validation cohort.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

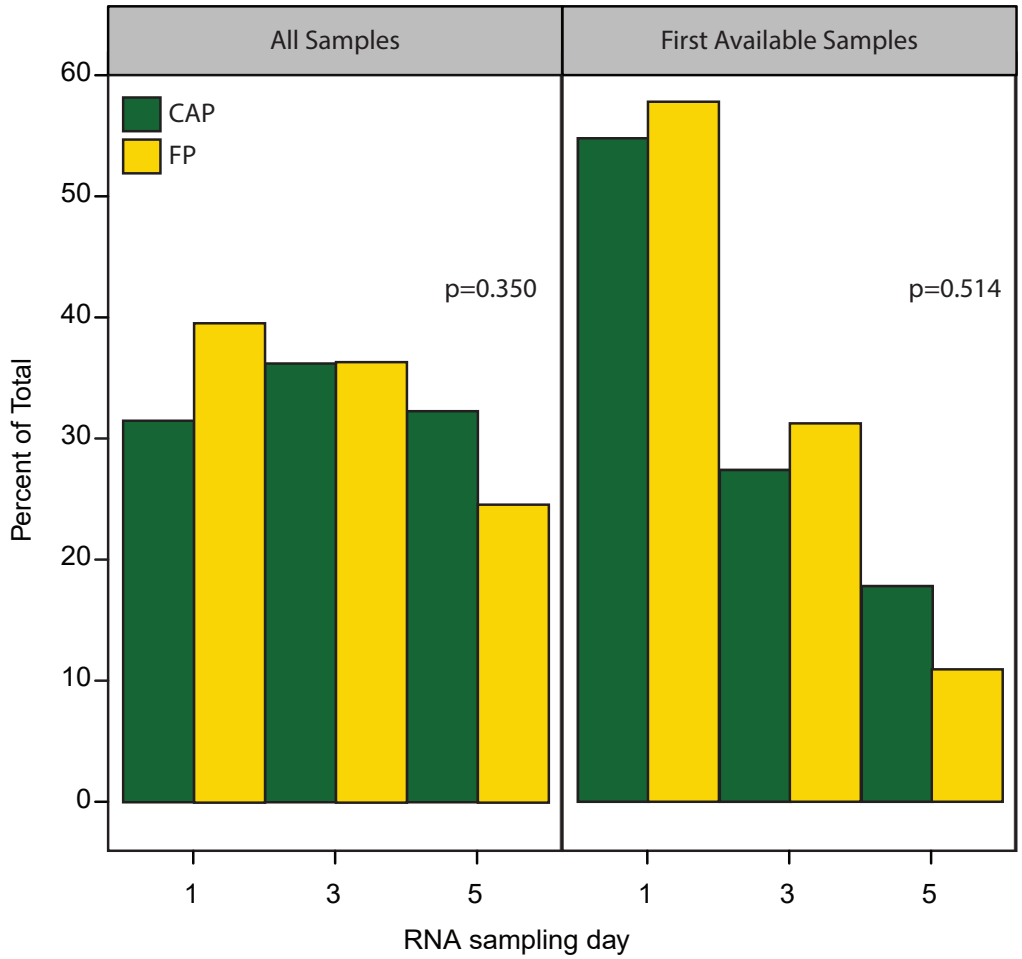


Figure E10. Histogram of RNA sampling day for CAP and FP samples.
The percentage of samples in the discovery cohort for each day following ICU admission are given for CAP (green) and FP (yellow). The left hand plot shows the percentage for all available serial samples, and the right hand plot shows the proportions when only the first available sample for each patient is selected. The CAP and FP subgroups do not differ significantly in terms of the proportion of samples from each day of sampling (chi-square test p-values shown).

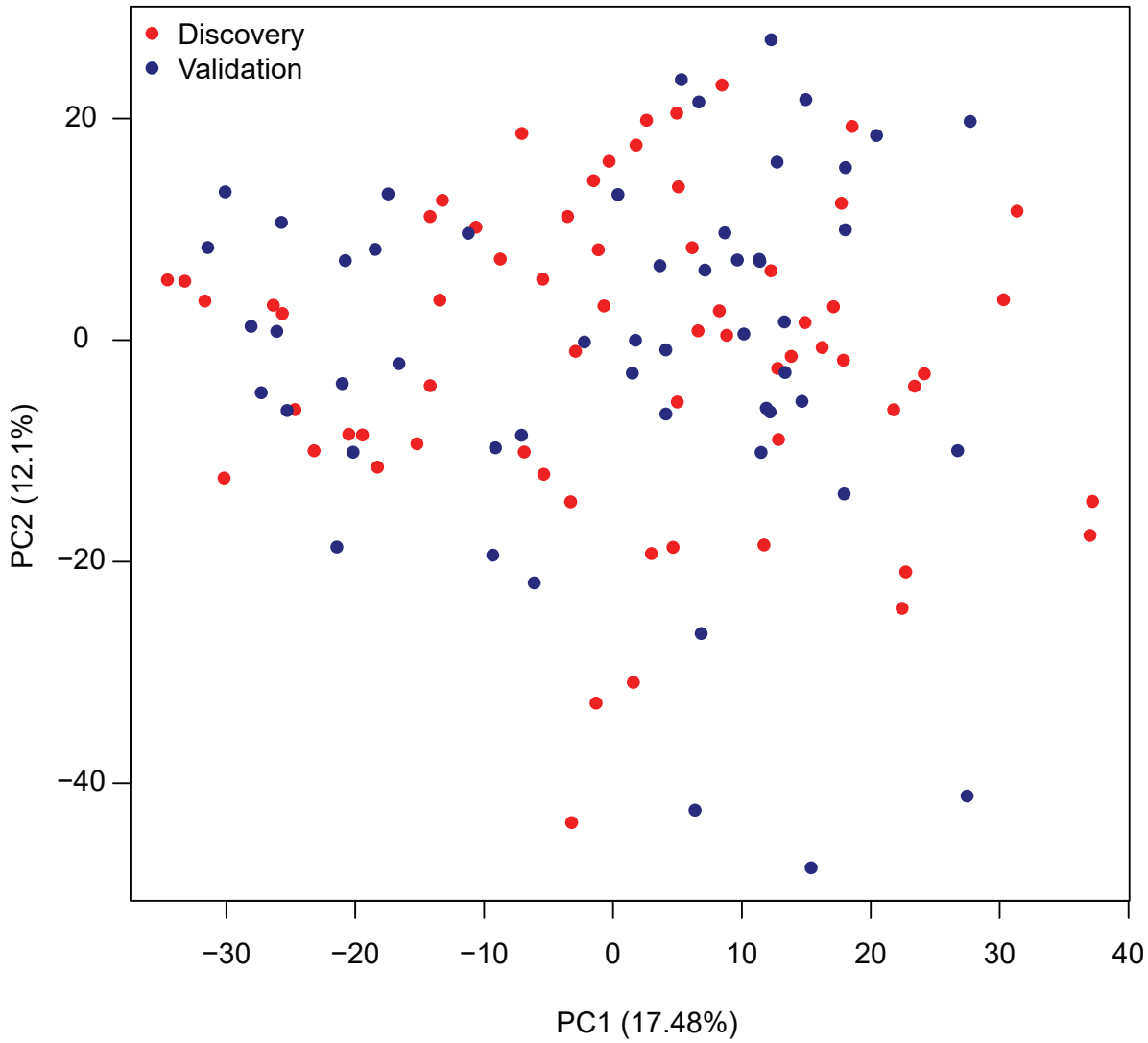


Figure E11. Principal component analysis (PCA) of FP samples from the discovery and validation cohorts combined using ComBat.
Following use of ComBat to account for batch effects between the discovery and validation cohorts, there was no clear segregation between the two sets of FP samples.

AJRCCM Articles in Press. Published on 30-December-2016 as 10.1164/rccm.201608-1685OC

Table E2: Participating hospitals involved in patient recruitment and GAinS Investigators.

Site	Principal Investigator	Research Nurses/Fellows	Other investigators
Aberdeen Royal Infirmary	Professor Nigel Webster	Jane Taylor, Sally Hall, Jenni Addison, Sian Roughton, Heather Tennant	
Broomfield Hospital, Chelmsford	Dr John Durcan/ Dr D Arawwawala	Karen Swan, Sarah Williams, Susan Smolen, Christine Mitchell-Inwang	
Charing Cross Hospital, London	Dr Tony Gordon	Emily Errington, Maie Templeton	
Hammersmith Hospital, London	Dr Stephen Brett	David Kitson, Robert Wilson, Laura Mountford, Juan Moreno	
John Radcliffe Hospital, Oxford	Dr Christopher Garrard/ Dr Stuart McKechnie	Paula Hutton, Penny Parsons, Alex Smith, Roser Faras-Arraya	
Norfolk & Norwich University Hospital, Norwich	Dr Simon Fletcher	Melissa Rosbergen, Georgina Glister	
Royal Berkshire Hospital, Reading	Dr Atul Kapila	Nicola Jacques, Jane Atkinson, Abby Brown, Heather Prowse	
Royal Victoria Infirmary, Newcastle	Dr Simon Baudouin	Charley Higham, Helen Walsh, Verity Calder, Catherine Swan, Heather Payne	
St Bartholomew's/Royal London Hospitals	Professor Charles Hinds	Eleanor McLees, Alice Purdy	Dr D. Watson
Southend Hospital	Dr David Higgins	Sarah Andrews, Sarah Mappleback	
University College London Hospital (UCLH), London	Dr Geoffrey Bellingan	Jung Hyun Ryu, Georgia Bercades, Susan Boluda	
Royal Blackburn Hospital	Dr Anton Krige	Martin Bland, Lynne Bullock, Donna Harrison	
Frenchay Hospital, Bristol	Dr Jasmeet Soar	Sally Grier, Elaine Hall	
Southmead Hospital, Bristol	Dr Jasmeet Soar	Sally Grier, Elaine Hall	
Royal Sussex County Hospital, Brighton	Dr Stephen Drage	Laura Ortiz-Ruiz De Gordoa, Sarah Lowes	
Huddersfield Royal Infirmary	Dr Peter Hall	Jackie Hewlett	
The James Cook University Hospital, Middlesbrough	Dr Stephen Bonner	Keith Hugill, Victoria Goodridge, Louise Cawthor	Dr Iain Whitehead
Kettering General Hospital	Dr Phil Watt	Parizade Raymode	
Leicester Royal Infirmary	Dr Jonathan Thompson	Sarah Bowrey, Sandra Kazembe, Natalie Rich, Prem Andreou, Dawn Hales, Emma Roberts	

Royal Preston Hospital, Preston	Dr Shond Laha	Jacqueline Baldwin, Angela Walsh, Nicola Doherty	
Queen Elizabeth Hospital, Birmingham	Professor Julian Bion	Joanne Millar, Elsa Jane Perry, Heather Willis, Natalie Mitchell, Sebastian Ruel, Ronald Carrera, Elsa Jane Perry, Jude Wilde, Annette Nillson, Sarah Lees	
St Mary's Hospital, London	Dr Martin Stotz	Adaeze Ochelli-Okpue	
Wythenshawe Hospital, Manchester	Dr Andrew Bentley	Katie Mccalman, Fiona Jefferies	
The Whittington Hospital, London	Dr Martin Kuper	Sheik Pahary	