Presymptomatic diagnosis of postoperative infection and sepsis using gene expression signatures

Roman A Lukaszewski[1,4], Helen E Jones[1], Vivian H Gersuk[2], Paul Russell[1,3], Andrew Simpson[1],

David Brealey[4,5], Jonathan Walker[6], Matt Thomas[7], Tony Whitehouse[8], Marlies Ostermann[9],

Alexander Koch[10,11], Kai Zacharowski[11], Mogens Kruhoffer[12], Damien Chaussabel[2,13], Mervyn Singer[4,5]

Correspondence: Prof Mervyn Singer, Bloomsbury Institute of Intensive Care Medicine, Division of Medicine, University College London, London UK.

email: m.singer@ucl.ac.uk ORCID 0000-0002-1042-6350

[1]Defence Science and Technology Laboratory, Porton Down, Salisbury, Wiltshire, UK.

[2]Benaroya Research Institute, Seattle, WA 98101-2795, USA.

[3]Salisbury NHS Foundation Trust, Salisbury, Wiltshire, UK.

[4]Bloomsbury Institute of Intensive Care Medicine, Division of Medicine, University College London, London UK.

[5]Division of Critical Care and, NIHR University College London Hospitals Biomedical Research Centre, University College London Hospitals NHS Foundation Trust, London, UK.

[6]Royal Liverpool & Broadgreen University Hospitals NHS Trust, Liverpool, UK.

[7]University Hospitals Bristol NHS Foundation Trust, Bristol, UK.

[8]University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Edgbaston, Birmingham, UK.

[9]Intensive Care Unit, Guy's and St Thomas's NHS Foundation Trust, London, UK

[10] Klinikum Esslingen, Esslingen 73707, Germany.

[11]Department of Anesthesiology, Intensive Care Medicine & Pain Therapy, University Hospital Frankfurt, Goethe University, 60590 Frankfurt am Main, Germany.

[12]BioXpedia A/S, DK-8200 Aarhus, Denmark.

[13]Laboratory of Translational Systems Immunology, Sidra Medicine, Doha, Qatar.

Abstract

Purpose: Early accurate diagnosis of infection ± organ dysfunction (sepsis) remain major challenges in clinical practice. Utilizing effective biomarkers to identify infection and impending organ dysfunction *before* the onset of clinical signs and symptoms would enable earlier investigation and intervention. To our knowledge, no prior study has specifically examined the possibility of pre-symptomatic detection of sepsis.

Methods: Blood samples and clinical/laboratory data were collected daily from 4,385 patients undergoing elective surgery. An adjudication panel identified 154 patients with definite postoperative infection, of whom 98 developed sepsis. Transcriptomic profiling and subsequent RT-qPCR was undertaken on sequential blood samples taken postoperatively from these patients in the three days prior to the onset of symptoms. Comparison was made against postoperative day-, age-, sex- and procedure- matched patients who had an uncomplicated recovery (n=151) or postoperative inflammation without infection (n=148).

Results: Specific gene signatures optimized to predict infection or sepsis in the three days prior to clinical presentation were identified in initial discovery cohorts. Subsequent classification using machine learning with cross-validation with separate patient cohorts and their matched controls gave high Area Under the Receiver Operator Curve (AUC) values. These allowed discrimination of infection from uncomplicated recovery (AUC 0.871), infectious from non-infectious systemic inflammation (0.897), sepsis from other postoperative presentations (0.843), and sepsis from uncomplicated infection (0.703).

Conclusion: Host biomarker signatures may be able to identify postoperative infection or sepsis up to three days in advance of clinical recognition. If validated in future studies, these signatures offer potential diagnostic utility for postoperative management of deteriorating or high-risk surgical patients and, potentially, other patient populations.

Trial Reg Number: ISRCTN17375399.   **Keywords:** Sepsis, diagnosis, host, biomarker, signatures

Take-Home Message

Early transcriptomic changes offer potential diagnostic utility for the management of patients at risk of developing subsequent postoperative infection ± sepsis. The limited number of genes identified facilitates development of a point-of-care rapid diagnostic.

140 character tweet

Large multicentre study predicting development of postoperative infection and sepsis up to 3 days before symptom onset

Statements and Declarations

RL- Dr. Lukaszewski reports grants from the US Defense Threat Reduction Agency, UK Ministry of Defence and UK Home Office during the conduct of the study. He has a patent Gene-based Methods and Associated Uses, Kits and System for Assessing Sepsis. GB2015002.5 Filed Sept. 2020 pending, a patent Gene-based Methods and Associated Uses, Kits and System for Assessing Sepsis. GB2015004.1 Filed Sept. 2020 pending, a patent Gene-based Methods and Associated Uses, Kits and System for Assessing Sepsis. GB2015007.4 Filed Sept. 2020 pending, a patent Sepsis Nucleic Acid Markers. GB1402293.3 Filed Jan 2014 issued, and a patent Protein-based Methods and Associated Uses, Kits and System for Assessing Sepsis. GB2009401.7 Filed June 2020, pending. He is a part-time advisor to PreSymptom Health Ltd who have licensed patents belonging to the UK Ministry of Defence.

HJ - Dr. Jones reports grants from US Defense Threat Reduction Agency, grants from UK Ministry of Defence, grants from UK Home Office, during the conduct of the study. In addition, Dr. Jones has a patent Sepsis Nucleic Acid Markers. GB1402293.3 Filed Jan 2014 issued.

MS - Prof Singer reports grants from the Defence Science and Technology Laboratory during the conduct of the study; grants from NewB, Apollo Therapeutics and UCL Technology Fund, other from

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by R Lukaszewski, H Jones, V Gersuk, D Brealey, J Walker, M Thomas, T Whitehouse, M Ostermann, A Koch, K Zacharowski, M Kruhoffer and D Chaussabel. The first draft of the manuscript was written by R Lukaszewski, M Kruhoffer and M Singer and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Introduction

Sepsis, the dysregulated host response to infection leading to life-threatening organ dysfunction [1], is a substantial global cause of mortality and morbidity [2]. Early, accurate diagnosis of infection and organ dysfunction remains problematic, as reflected by multiple interventional trials over three decades failing to yield outcome improvements [3]. These failures relate in part to uncertainty that infection was actually present [4, 5], but also to belated intervention once the patient is in established multi-organ failure with the therapeutic window of opportunity having closed [3, 6]. Success with immunomodulatory and other therapeutic strategies is likely predicated on use early in the sepsis course or even pre-emptively, as frequently achieved in preclinical models [6]. On the other hand, inappropriate use of antibiotics contributes to antimicrobial resistance [*7]* and may distract clinicians from diagnosing a non-infective condition [4, 5].

Considerable efforts are being expended to develop rapid, even point-of-care, biomarkers for infection and sepsis with high sensitivity and specificity [8]. Pathogen-focused techniques include molecular identification of the pathogen or unique components such as prokaryotic DNA. The dysregulated host response is being interrogated using -omic approaches, single- or multiplex protein assays or flow cytometry [8, 9]. However, host response studies to date often lack comparator cohorts of patients with non-infective causes of critical illness, or microbiological confirmation that infection is indeed present. Furthermore, sampling is usually begun *after* the patient presents with suspected infection or sepsis. An ideal patient cohort would allow accurate pre-symptomatic identification of patients who proceed to develop infection, in particular those progressing to sepsis.

Our primary study aim was to develop a tool to facilitate presymptomatic identification of patients developing infection, and we hypothesised this could be achieved using transcriptional changes in small gene sets. As a secondary objective, we sought to determine whether we could pre-identify those infected patients who would develop new-onset organ dysfunction (sepsis). To this end, we

6

conducted a large prospective, multi-centre study in patients undergoing elective major surgery, with daily blood sampling and data recording commencing preoperatively and continuing up to a week after. A clinical adjudication panel independently examined clinical and laboratory data to identify patients with definite infection ± sepsis. Samples from these patients enabled comparison of microarray and RT-qPCR data against cohorts of postoperative day-, age-, sex- and procedure-matched patients with either non-infective systemic inflammation or an uncomplicated postoperative course.

MATERIALS AND METHODS

Study design

Elective surgery patients were prospectively recruited at eight hospitals (seven UK, one German) between November 2007 and February 2017. Ethical approval was granted through the Southampton and South-West Hampshire Multicentre Research Ethics Committee (Reference 06/Q1702/152). The protocol achieved US Federal Wide Assurance Independent Review Board status (IRB00001756). Patients were enrolled if they gave informed consent, were aged between 18-80 years, and due to undergo elective major surgery that would likely enhance the development of postoperative infection (Supplement S1).

Data and blood sampling collection

Demographic data were collected at enrolment with baseline blood sample collection taken at 1-7 days pre-operatively. After surgery, relevant clinical, laboratory and imaging data and blood samples were collected daily on all patients until seven days, hospital discharge (if sooner), or a diagnosis of infection or sepsis by the treating clinician. Two 4 ml aliquots of blood were collected daily into sterile EDTA vacutainers and then immediately transferred into RNAse-free vials containing 10.5 ml RNA*later*® (ThermoFisher, Waltham, MA, USA) to preserve the transcriptome.

Patient selection process

A detailed description is provided in Supplement S2. Briefly, initial diagnosis of sepsis was based on the treating clinician's interpretation of clinical and laboratory markers using the then-extant 'Sepsis-2' definition of sepsis [10]. This described '*sepsis*' as suspected or confirmed infection with two or more systemic inflammatory response syndrome (SIRS) criteria, and '*severe sepsis*' as sepsis plus poorly-characterized new-onset organ dysfunction. The new sepsis definition (Sepsis-3), published in February 2016 redefined '*sepsis*' as infection plus new organ dysfunction identified by a ≥2 point rise in the Sequential Organ Dysfunction Assessment (SOFA) score [1]. To align with modern nomenclature, subsequent analyses and descriptors apply the new definition.

Since clinical opinion can vary when diagnosing infection or sepsis [11], a Clinical Advisory Panel adjudicated cases and enabled confident identification of patients with postoperative infection (Supplement S2). Patients with definite infection were allocated to either an uncomplicated infection subgroup, or to a sepsis subgroup if new organ dysfunction developed following surgery with a rise in Sequential Organ Failure Assessment (SOFA) score ≥2 points above baseline [1]. Postoperative day-, age-, sex- and procedure-matched cohorts of non-infected patients making an uncomplicated postoperative recovery (SIRS-), or developing a systemic inflammatory response (SIRS+), were selected from the remaining pool of patients recruited into the study. Blood samples from these three groups, totalling 453 patients, were taken forward for transcriptomic and subsequent reverse transcription quantitative polymerase chain reaction (RT-qPCR) analyses.

Microarray and RT-qPCR analysis

Infected patients were split into Discovery (Gene Feature Selection) (n= 63) and Training and Validation (Classification) (n=91) cohorts (Figure 1).

• *Discovery (Gene Feature Selection) Cohort*

Biomarker discovery studies were undertaken exclusively in this 63-patient cohort at a pre-determined interim analysis of the study. A detailed transcriptomic analysis was initially undertaken using globin-reduced RNA (GlobinClear Human, ThermoFisher) from the blood samples of 58 patients collected during the three days prior to presentation of definite infection, and from 55 matched non-infected (SIRS-) control patients (Figure 1). Further details of the microarray analysis using Human HT-12v4 beadarrays (Illumina, San Diego, CA, USA) are shown in Supplement S3. Differentially expressed genes (DEGs) based on microarray data were identified by applying a linear model fit for each gene [12], a p-value cut-off of 0.05 and a fold change of at least 1.2. Data obtained from non-infected study patients were used as reference. Gene expression of the top 80 identified DEGs was verified using multiplex RT-qPCR (Fluidigm, San Francisco, CA, USA) and TaqMan Gene Expression Assays (Applied Biosystems, Carlsbad, CA, USA).

The RT-qPCR data were then further used for downselecting the 25 most promising biomarkers. In these experiments, analysis was possible for 63 infection and 62 SIRS- patients. In addition, RT-qPCR data were obtained from 58 SIRS+ matched patients (Figure 1, Table 1). Within the 63 infected patients, 37 developing sepsis were compared to 26 with uncomplicated infection (Figure 1, Table 1). Gene ontology enrichment was performed using the R package clusterProfiler (v3.18.0) and visualized using the R package GoPlot (v1.0.2).

Down-selection of genes based on RT-qPCR data was made using a two-step feature selection process to improve the performance of the predictive classification models used to discern patient groups (described fully in Supplement S4). Briefly, the Boruta algorithm [13], a wrapper method based on random forest [14] was used for selection of relevant features in the dataset. This step was followed by backward elimination to determine features with the most discriminatory power for a particular classification. Training of these models relevant for feature selection was done by a 5-fold cross-validation repeated 25 times. Mean performances were tracked for AUC, PPV and NPV over the 25 repetitions. This procedure was repeated to assess the ability of the identified biomarkers to discriminate between infected patients and either controls (SIRS-) or non-infected but inflamed (SIRS+) patients and, as a secondary objective, to discriminate between infected patients either developing organ dysfunction (sepsis) or not (Figure 1).

• *Training and validation of final classification models*

Of all 80 DEG biomarker candidates, 25 were identified as part of at least one classification model based on an orthogonal RT-qPCR approach that could discriminate between the patient classes. These 25 RT-qPCR transcripts were taken forward to an independent patient cohort (Classification Cohort) to both learn a predictive classifier model and to assess performance.

RT-qPCR analysis of gene expression was undertaken on samples from 91 patients with infection, 89 matched SIRS- comparator patients, and 90 similarly matched non-infected SIRS+ patients, respectively (Figure 1). Further comparison was made within the infection cohort, comparing patients who did (sepsis, n=61) or did not develop organ dysfunction (n=30). A final comparison was made between patients with sepsis (n=61) and all other matched patients with non-sepsis outcomes (n=209) within our validation cohorts (Figure 1).

For assessment of signature performance, a similar random forest-based approach was used. Models were created ('training phase') and and then validated ('validation phase'). We divided our data in a random 80:20 split for training and validation, respectively. In the training phase models were trained and cross-validated on the 80% subset using 5-fold cross-validation. This process was repeated 25 times using R package caret (v6.0-90, 'train' function) to prevent lucky or unlucky 5-fold cross-validation. In the validation phase the trained models were tested on the remaining 20% of the data that were not used for cross-validation in the training phase. This process was repeated 10 times with 10 different randomly selected subsets, again to counteract any lucky or unlucky data splits. For each comparison, 10 models (including 10 different test sets) were created and the mean AUC, sensitivity and specificity reported.

Statistical analysis

Prediction of a robust sample size followed the method described by Figueroa *et al* [15] and is dscribed in detail in Supplement S4. Briefly, random forest-based classifier performance was assessed with an increasing number of simulated patients in each cohort to identify when a performance plateau occurred.

All qRT-PCR data was analyzed with nonparametric Wilcoxon tests, unless otherwise stated. All data and statistical analyses were done using R (v.3.6.3).

RESULTS

Patient recruitment and subgroup identification

Clinical data and blood samples were collected from 4,385 high-risk elective surgery patients. An overall flow chart is shown in Supplement S5. The clinical panel adjudication identified a total of 154 patients with high certainty of infection, of whom 98 developed sepsis (Figure 1). These were compared against samples taken from postoperative day-, age-, sex- and operation-matched patients without infection, of whom 148 developed a systemic inflammatory response (SIRS+) and 151 (SIRS-) did not. Each patient group was split into Discovery (Gene Feature selection) and Training/Validation (Classification) cohorts. Demographic and clinical data are shown in Table 1. Microbiological confirmation of infection was achieved in 81 septic and 45 uncomplicated infection patients (Supplement S6). A broad range of Gram negative, Gram positive and fungal organisms were isolated. Strong clinical evidence of infection was used for the remaining 15 and 11 patients, respectively. Of the 98 septic patients, 80 patients required at least one organ support (60 mechanical ventilation, 55 vasopressors, 8 renal replacement therapy) Nine patients ultimately died of sepsis while two died for unrelated reasons. Approximately half of the patients who were subsequently diagnosed with infection and 60% of the non-infected patients received no postoperative antibiotic (Supplement S7). The remaining patients received one or more doses of antibiotic, 79% commencing on Postoperative Day 1 and likely to represent ongoing prophylactic medication.

Discovery (Gene Feature selection) cohort

• *Microarray-based biomarker discovery*

Initial biomarker discovery was performed on blood samples taken from 58 infected patients over the three days preceding clinical presentation of postoperative infection underwent transcriptomic profiling. Comparison was made against 55 postoperative day-, age-, sex- and procedure-matched patients with an uncomplicated, non-infected, non-inflamed (SIRS-) course (Figure 1, Table 1).

Overall, 2,594 differentially expressed genes (DEGs) between infection and control were identified in the three days pre-infection diagnosis. Of the top 1500 DEGs with the highest fold change (Figure 2A, Supplement S8), 863 (57.5%) DEGs were upregulated and 637 (42.5%) downregulated. Functional enrichment analysis yielded immune relevant pathways involving primarily neutrophils and T cells (Figure 2B, Supplement S8).

A random forest-based algorithm identified the best-performing DEGs derived for all three days prior to infection/sepsis diagnosis and their respective non-infected controls. Initial analysis indicated that different sets of eight genes could identify postoperative infection with high AUCs.

• RT-qPCR verification and down-selection of microarray-based biomarkers

Collectively, differential expression analysis and concurrent random forest-based classifiers with equally high performance metrics proposed 80 best performing genes from the microarray data. These were down-selected for confirmation by RT-qPCR analysis using the same 58 infected patient cohort samples analysed for biomarker discovery plus samples from an additional five patients, with comparisons being made against 62 non-infected SIRS- patients. A representative 8-gene-set showed highly significant differential expression between the two patient groups, exemplifying the compatibility between microarray- and RT-qPCR-based gene expression (FDR-corrected Wilcoxon test, p<0.001, Fig. 2C). To utilise all patient sub-groups with different postoperative outcomes identified in our study, we subsequently explored whether subsets of our 80-gene set would suffice to differentiate the 63 infected patients from 58 SIRS+ patients, as well as to differentiate within the infection cohort, i.e. with (n=37) or without (n=26) organ dysfunction (Figure 1, Table 1). Again using a random forest-based classification approach in conjunction with Boruta to select the most important gene transcripts (*cf* Methods) yielded convincing model performances throughout all classification attempts. Specifically, a 7 gene set (*B4GALT5, AFF1, LDLR, ATXN7L3, LARP4B, SLC36A1, TRPM2*, AUC >0.85, PPV >0.8) for infection versus SIRS- models, a 12 gene set (*ATXN1, SLC41A3, MED13L, STOM, B4GALT5, MIDN, HVCN1, LDLR, CFLAR, SPATA13, EIF4G3, METTL7B*, AUC >0.9, PPV >0.8) for infection versus SIRS+ models, and an eight gene set (*DOK3, ICAM2, IL1R1, LGALS2, LSG1,*

RPL13A, RPS13, SGSH, AUC >0.75 (PPV >0.7) for sepsis vs. uncomplicated infection were sufficient to classify postoperative outcomes. These 25 genes were further selected for building predictive models in an independent and substantially larger patient cohort.

Training/Validation (Classification) Cohort:

• Building machine learning predictive models in independent patient cohorts

In separate cohorts of patients we evaluated the classification performances of the 7, 12 and 8 gene signatures whose performance had been verified to discern: (i) infection (n=91) from non-infected SIRS- (n=89) patients, (ii) infection from non-infected SIRS+ (n=90) patients, and (iii) 61 sepsis and 30 uncomplicated infection patients, respectively, (Figure 1, Table 1).

To prevent models from overfitting, classification training including nested cross-validation was performed 10 times on a randomly selected 80% of the data. In each iteration the models were validated on 20% of randomly selected sample subsets that were never used for training (kept outside of cross-validation, *cf.* Methods). The mean performance of the gene signatures across these 10 runs remained high following a random forest-based classification. For postoperative infection versus SIRS- controls, the 7 gene signature achieved an AUC value of 0.871 (Figure 3). The 12 gene signature for differentiation of infection from non-infected SIRS+ patients achieved an AUC value of 0.897. Differentiation of sepsis from uncomplicated infection patients resulted in an AUC value of 0.703. Finally, differentiation of sepsis from all other clinical presentations using all 25 transcripts achieved an AUC of 0.843. Sensitivity was high across all comparisons (0.785-0.942) whereas specificity was high for infected compared to non-infected SIRS+ (0.838) and SIRS- (0.776) patients, but poor for sepsis vs uncomplicated infection (0.217). Supplement S9 shows the Classification performance for different thresholds optimized for sensitivity or specificity, and various assumed prevalence of outcomes. This indicates, for example, how current practice specificity for correct antibiotic selection (Supplement 7) can be enhanced by use of a specificity-optimized gene panel.

DISCUSSION

This study demonstrates that patients undergoing elective major surgery who develop postoperative infection with or without organ dysfunction (sepsis), can be reliably identified and differentiated from non-infected patients up to three days before clinical diagnosis using host transcriptomics.

Accurate, early diagnosis of infection and sepsis represents a holy grail for improving patient care and outcomes and avoiding unnecessary antimicrobial use. Transcriptomic studies of patients with uncomplicated infection and/or sepsis have focused on patients with established clinical features requiring hospitalization or admission to critical care [17-21]. Such studies have added to our understanding of sepsis pathogenesis, demonstrating some prognostic capability [18, 22-24], and offering reasonable discrimination between infectious and non-infectious causes of systemic inflammation [18-20, 25]. Presymptomatic diagnosis would facilitate early investigation and treatment, and can be reasonably surmised to improve patient outcomes.

To our knowledge, no study has specifically studied patient samples taken *before* the clinical onset of infection. An elective surgical population represents an ideal patient group, albeit somewhat inefficient, as most have a postoperative course uncomplicated by infection. Apart from the underlying need for surgery, patients enrolled in our study were infection-free and stable. To have as clean a dataset as possible, samples were only used from the 3.5% of patients where a clinical adjudication panel identified postoperative infection with high confidence. Intermediate cases with more uncertainty surrounding diagnosis were excluded. We also drew from a large patient cohort from which postoperative day-, age-, sex- and operation-matched patients could be selected having either an uncomplicated course or developing a postoperative systemic inflammation where infection was confidently excluded and not treated.

Many studies examining sepsis biomarkers have relied on selection through knowledge-based approaches predicated on known biological functions and pathways [25] rather than targets whose

15

functions have not been fully described [26]. We adopted a target selection strategy that relied on statistical rather than biological features [27]. A machine learning approach was used to select appropriate targets and classify patients based on host gene expression with comparison against a clear-cut clinical diagnosis to determine predictive accuracy. The initial phase of biomarker discovery identified 80 genes for which differences from other patient cohorts could be confirmed using an orthogonal approach (RT-qPCR). Re-classification based on RT-qPCR aimed at feature reduction for best gene selection in the same group of patients enabled further down-selection to 25 key biomarkers. Subsequent classification analysis revealed the potential for clinical application with an even more reduced set of biomarkers, allowing a rapid and affordable point-of-care test..

A machine learning algorithm approach was also used to select appropriate targets and classify patients based on host gene expression (uncorrected for baseline values) with comparison made against a clear-cut clinical diagnosis to determine predictive accuracy. The success of this approach is evidenced by high values of AUC, sensitivity and specificity comparing patients developing infection versus uncomplicated (SIRS-) controls, and the more diagnostically challenging group who develop postoperative (SIRS+) inflammation that is not driven by infection yet which shares many clinical features such as pyrexia, tachycardia, neutrophilia and elevated C-reactive protein levels. There was also reasonable discrimination (high sensitivity but poor specificity) in pre-emptively identifying infected patients who would proceed to organ dysfunction compared to those with uncomplicated infection.

The difference in our study approach is also exemplified by only six out of 80 genes identified in our classification models overlapping with genes of prior published gene signatures for sepsis and community-acquired pneumonia [16, 18, 20]. The data reported within our study indicate that changes in expression of host biomarkers are predictive of later complications related to infection. Any shortcomings in model performance is likely due to the complex relationship between individual genes, their transcripts, post-translational modifications and protein interactions, and complicated

further by both therapeutic interventions and the dynamic nature of signaling within the host response that flux during the dysregulation process [27, 28].

In trying to answer the more challenging question of when an individual will develop infection or sepsis, we sampled at multiple timepoints. While a common gene set could detect changes over a 3-day period prior to clinical presentation, separate gene transcripts gave even stronger signals on a day-by-day basis (data not shown). This reflects observations made in trauma patients [29], in human volunteers given endotoxin [30], and in children with meningococcal disease [31]. However, using separate gene sets would require prior knowledge of precisely when a patient will develop features of uncomplicated infection or sepsis, and this would clearly be non-feasible in clinical practice.

Several limitations should be highlighted. Cost issues precluded even broader testing, in particular within the indeterminate subgroup in whom there was less certainty about the presence of postoperative infection but where antibiotics were commenced. We selected 20% patient subsets for internal model validation, which we repeated randomly 10 times to avoid a single selected advantageous or disadvantageous sample subset. Prospective external validation is needed to confirm predictive utility in both surgical and non-surgical populations, in specific patient subsets (e.g. children, haemoncology), in a broader racially diverse mix, and in patients with specific non-bacterial infections such as malaria and SARS-CoV-2. A validation study in patients with acute lower respiratory tract infection is currently in progress.

How such gene panels, if validated as reliable tools, could be used in clinical practice will depend on both cost and the rapidity with which results can be obtained. The ability already exists to deliver point-of-care information from small gene set PCR-based devices within 1-2 hours, or from multiplex protein analyses within minutes. This can currently be achieved at relatively low cost and the price per unit test is likely to drop further with advances in technology and manufacturing. If cost is comparable to routinely monitored biomarkers of inflammation such as C-reactive protein and

procalcitonin, such a test could potentially offer (near-) daily 'rule-out' screening. The target product profile could also be altered using different cut-off values (as shown in Supplement S9) to enhance either specificity. Especially if the cost-economic analysis mitigates against its use as a daily monitor, it could still be used as a diagnostic with rule-in capability focussing on patients in whom infection is clinically suspected but where diagnosis remains uncertain. A further potential use is a a 'rule out' test in patients presenting with non-specific features where infection is among a list of differential diagnoses. And, finally, the high sensitivity achieved from the complementary gene set in discriminating future uncomplicated infection from sepsis offers potential for early, safe discharge for low-risk patients from a higher dependency ward setting. Modelled examples comparing optimized sensitivities and specificities are shown in Supplement 10.

In conclusion, we demonstrate that postoperative infection can be identified by a panel of gene transcripts up to three days pre-onset of clinical presentation. This can be delineated from both uncomplicated postoperative controls and, more crucially, patients with postoperative systemic inflammation unrelated to infection. Infected patients developing organ dysfunction (sepsis) could also be identified in advance. Subject to prospective confirmation, the identified gene sets offer the ability to reliably diagnose infection and sepsis presymptomatically, which may impact upon patient outcomes and inappropriate antibiotic use.

Acknowledgements

Data sharing

Cohort data are available in Table 1 and the Supplement. Microarray data are stored on Mendeley (doi: 10.17632/rhc5s6zj88.1) and will be available on reasonable request.

References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al (2016) The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 315: 801-810. https://doi.org/10.1001/jama.2016.0287.

2. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al (2020) Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. Lancet 395: 200–11. https://doi.org/ 10.1016/S0140-6736(19)32989-7.

3. Vincent J-L, Sakr Y (2019) Clinical trial design for unmet clinical needs: a spotlight on sepsis. Expert Rev Clin Pharmacol 12: 893–900. https://doi.org/ 10.1080/17512433.2019.1643235.

4. Heffner AC, Horton JM, Marchick MR, Jones AE (2010) Etiology of illness in patients with severe sepsis admitted to the hospital from the emergency department. Clin Infect Dis 50: 814–20. https://doi.org/10.1086/650580.

5. Klein Klouwenberg PMC, Cremer OL, van Vught LA, Ong DS, Frencken JF, Schultz MJ, et al (2015) Likelihood of infection in patients with presumed sepsis at the time of intensive care unit admission: a cohort study. Crit Care 19:319 https://doi.org/10.1186/s13054-015-1035-1.

6. Perner A, Gordon AC, Angus DC, Lamontagne F, Machado F, Russell JA, et al (2017) The intensive care medicine research agenda on septic shock. Intensive Care Med 43: 1294–305. https://doi.org/10.1007/s00134-017-4821-1.

7. World Health Organisation: Antimicrobial resistance, https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance (Accessed 19 November 2021).

8. Opal SM, Wittebole X (2020) Biomarkers of infection and sepsis. Crit Care Clin 36: 11–22. https://doi.org/10.1016/j.ccc.2019.08.002.

9. Kim M-H, Choi J-H (2020) An update on sepsis biomarkers. Infect Chemother 52: 1-18. https://doi.org/10.3947/ic.2020.52.1.1.

10. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cohen J, Opal SM, Vincent JL, Ramsay G (2003) 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Crit Care Med 31: 1250–6. https://doi.org/ 10.1097/01.CCM.0000050454.01978.3B.

11. Rhee C, Kadri SS, Danner RL, Suffredini AF, Massaro AF, Kitch BT, et al (2016) Diagnosing sepsis is subjective and highly variable: a survey of intensivists using case vignettes. Crit Care 20: 89. https://doi.org/10.1186/s13054-016-1266-9.

12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43: e47–e47. https://doi.org/10.1093/nar/gkv007.

13. Kursa MB, Rudnicki WR (2010) Feature selection with the Boruta Package. J Stat Soft 36. https://doi.org/10.18637/jss.v036.i11.

14. Breiman L (2001) Random Forest. Machine Learning 45: 5–32.

15. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH (2012) Predicting sample size required for classification performance. BMC Med Inform Decis Mak 12: 8. https://doi.org/10.1186/1472-6947-12-8.

16. Cazalis M-A, Lepape A, Venet F, Fragar F, Mougin B, Vallin H, et al (2014) Early and dynamic changes in gene expression in septic shock patients: a genome-wide approach. Intensive Care Med Exp 2: 20. https://doi.org/10.1186/s40635-014-0020-3.

17. Langley RJ, Tsalik EL, Velkinburgh JC, Glickman SW, Rice BJ, Wang C, et al (2013) An integrated clinico-metabolomic model improves prediction of death in sepsis. Sci Transl Med 5: 195ra95–195ra95. https://doi.org/10.1126/scitranslmed.3005893.

18. Sweeney TE, Shidham A, Wong HR, Khatri P (2015) A comprehensive time-course–based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. Sci Transl Med 7: 287ra71–287ra71. https://doi.org/10.1126/scitranslmed.aaa5993.

19. Miller RR, Lopansri BK, Burke JP, Levy M, Opal S, Rothman RE, et al (2018) Validation of a host response assay, SeptiCyte LAB, for discriminating sepsis from systemic inflammatory response syndrome in the ICU. Am J Resp Crit Care Med 198: 903–13. https://doi.org/10.1164/rccm.201712-2472OC.

20. Scicluna BP, Klein Klouwenberg PMC, van Vught LA, Wiewel MA, Ong DS, Zwinderman AH, et al (2015). A molecular biomarker to diagnose community-acquired pneumonia on intensive care unit admission. Am J Resp Crit Care Med 192: 826–35. https://doi.org/10.1164/rccm.201502-0355OC.

21. Burnham KL, Davenport EE, Radhakrishnan J, Humburg P, Gordon AC, Hutton P, et al (2017) Shared and distinct aspects of the sepsis transcriptomic response to fecal peritonitis and pneumonia. Am J Resp Crit Care Med 196: 328–39. https://doi.org/10.1164/rccm.201608-1685OC.

22. Sweeney TE, Perumal TM, Henao R, Nichols M, Howrylak JA, Choi AM, et al (2018) A community approach to mortality prediction in sepsis via gene expression analysis. Nature Commun 9: 694. https://doi.org/10.1038/s41467-018-03078-2.

23. Scicluna BP, van Vught LA, Zwinderman AH, Wiewel MA, Davenport EE, Burnham KL, et al (2017) Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. Lancet Resp Med 5: 816–26. https://doi/10.1016/S2213-2600(17)30294-1.

24. Pierrakos C, Vincent J-L (2010) Sepsis biomarkers: a review. Crit Care 14: R15. https://doi.org/ 10.1186/cc8872.

25. Sutherland A, Thomas M, Brandon RA, Brandon RB, Lipman J, Tang B, et al (2011) Development and validation of a novel molecular biomarker diagnostic test for the early detection of sepsis. Crit Care 15: R149. https://doi.org/10.1186/cc10274.

26. Kaplan JM, Wong HR (2011) Biomarker discovery and development in pediatric critical care medicine. Pediatr Crit Care Med 12: 165–73. https://doi.org/10.1097/PCC.0b013e3181e28876.

27. Maslove DM, Wong HR (2014) Gene expression profiling in sepsis: timing, tissue, and translational considerations. Trends Mol Med 20: 204–13. https://doi.org/10.1016/j.molmed.2014.01.006.

28. Tang BMP, McLean AS, Dawes IW, Huang SJ, Lin RCY (2009) Gene-expression profiling of peripheral blood mononuclear cells in sepsis. Crit Care Med 37: 882–8. https://doi.org/10.1097/CCM.0b013e31819b52fd.

29. Xiao W, Mindrinos MN, Seok J, Cuschieri J, Cuenca AG, Gao H, et al (2011) A genomic storm in critically injured humans. J Exp Med 208: 2581–90. https://doi.org/10.1084/jem.20111354.

30. Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, et al (2005) A network-based analysis of systemic inflammation in humans. Nature 2005; 437: 1032–7. https://doi.org/ 10.1038/nature03985.

31. Kwan A, Hubank M, Rashid A, Klein N, Peters MJ (2013) Transcriptional instability during evolving sepsis may limit biomarker based risk stratification. PLoS ONE 8: e60501. https://doi.org/10.1371/journal.pone.0060501.

Figure Legends

Figure 1: Study schema with Discovery (Gene Feature Selection) and Training/Validation (Classification) phases indicated.

RT-qPCR = Reverse Transcription quantitative polymerase chain reaction; SIRS = systemic inflammatory response syndrome; Inf+ = infected patients; Inf- = non-infected patients; Inf+ OD- =uncomplicated infection patients.

Figure 2: Gene expression analysis for *infection* versus *healthy postoperative controls*.

A) Expression profile of 1500 differentially expressed genes with highest absolute fold change. The top ten up- and downregulated genes are indicated separately. B) Top enriched Gene Ontology categories based on 1500 differentially expressed genes with highest fold change as shown in figure 2A). Node sizes indicate number of genes per shown category. Fold change is indicated for each displayed gene node. C) RT-qPCR values for 8 genes originating from classifying *infection* versus *no infection* based on microarray based expression data.

ΔCq value indicates PCR cycle quantifications according to ΔCq = Cq(reference) - Cq(target gene). Wilcoxon test derived significance level is indicated (***: p≤0.001, ns: not significant). Error bars show standard error of the mean. N(samples of infected patients) = 139, N(samples of control patients) = 144.

Figure 3: RT-qPCR based classification performance based on random forest using all available patient samples and across all days prior to infection diagnosis. ROC curve and statistical metrics for (i) infection versus non-inflamed postoperative controls (SIRS-) [blue], infection versus non-infected and inflamed (SIRS+) [red], sepsis versus infection without organ dysfunction (OD-) [mauve] and sepsis versus all others [green]. The table reports mean statistics for classification models based on different numbers of transcripts.

N = number; AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value; Sens = sensitivity; Spec = specficity

Table Legend

Table 1 Patient demographics and clinical metadata. Three main cohorts were used for analysis (infected, non-infected SIRS+ and non-infected SIRS- patients. The infection cohort is further sub-divided into patients who developed organ dysfunction (sepsis) and those that did not. Data are given for cohorts used for initial biomarker discovery and subsequent biomarker validation.

IQR interquartile range; SOFA Sequential (Sepsis-related) Organ Failure Assessment Score.