nature research

Corresponding author(s): Timothy A. Blauwkamp

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☐ | ☒ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☐ | ☒ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |
| ☐ | ☒ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No software was used. |
| Data analysis | The core software used as part of the Karius test is described in the "Clinical-grade Microbial cfDNA Sequencing for Infectious Disease" portion of the methods section, under sub-sections "Sequence data processing and alignment", "Microorganism abundances estimation", and "Pathogen detection". The open source software includes the following external tools: bcl2fastq v2.17.1.14, Trimmomatic v 0.32, Bowtie v2.2.4, and BLAST v2.2.30. A description of all open source code is included in the methods, and further details are available upon request. The proprietary portions of the code are not available. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon request. Sequencing data that support the finding of this study (with human reads removed) have been deposited in NCBI SRA and can be accessed with [DOIs to be provided when complete].

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | For all experiments, the sample size was informed by published CLSI guidelines for analytical validation experiments, peer-reviewed publications in related fields, and availability of clinical study samples. |
| Data exclusions | Criteria for data exclusion were pre-established and based on several quality control indicators. Any sample failing to meet quality control specifications for sequencing depth, sample-carryover limits, sample mix-ups or any sample contained within a batch where the positive or negative batch controls failed to meet the requirements for quality and accuracy were repeated. Any sample failing to meet quality control requirements upon repeat testing was excluded. |
| Replication | Indications of reproducibility are included with each experiment. The only experiments in which results were measured only once are those involving clinical study specimens. Even here, an indication of the reproducibility in testing of clinical samples was determined on twenty of these samples, as described in the precision testing section. The ability to replicate results obtained on other clinical samples is limited by the volume of sample available for repeated testing. |
| Randomization | Collection of plasma from self-reported healthy volunteer for the reference interval study was performed by a third party based on the criteria outlined in the materials and methods. Collection of clinical samples for comparison of testing to CMV-PCR was performed by a third-party with instructions to collect samples containing a full range of CMV infection levels, from very low to very high, as determined by CMV-PCR. The inclusion criteria for the 350 prospectively enrolled patients with clinical suspicion of sepsis are described in the methods section. |
| Blinding | The testing process is automated from sample preparation (performed on high throughput liquid handlers) through data analysis (performed by locked and version-controlled analysis pipelines), as are the criteria by which a test result is either accepted into the dataset or not. There is no human judgment involved in generating or accepting a test result. Nonetheless, all experimentation was performed while blinded to the results of blood culture and CMV-PCR testing for all clinical sample testing. |

# Reporting for specific materials, systems and methods

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Unique biological materials |
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☐ | ☒ Human research participants |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Unique biological materials

Policy information about availability of materials

| | |
|---|---|
| Obtaining unique materials | All materials used in this study are available from the standard commercial sources indicated in the manuscript, except for the clinical study samples. Due to extremely limited sample volumes for these clinical study samples, we are unable to make these available. |

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | Samples utilized in the determination of the reference range were collected from 167 healthy asymptomatic donors in geographically diverse areas of the US, aged 18 to 65 years of age, who had been screened for common health conditions including infectious diseases via a questionnaire and standard blood donor screening assays. Samples utilized in the quantitative evaluation of the assay included plasma from 25 CMV positive individuals obtained from a biorepository of de-identified remnant blood samples for which age and sex of the donors was collected. Samples from patients within the SEP-SEQ cohort were aged 18 years or older with suspected sepsis evidenced by a temperature of >38℃ or <36℃ and at least one additional Systemic Inflammatory Response Syndrome (SIRS) criteria. The first 1500 samples (~50% pediatric) tested sequentially by the Karius commercial laboratory were most commonly from immunocompromised patients, followed by patient suspected of sepsis, endocarditis, and complicated pneumonia. |
| Recruitment | Samples utilized for the reference range were obtained from specimen procurement companies who recruited screened healthy asymptomatic adult volunteers presenting to geographically diverse regional U.S. blood centers for routine plasma donation. Patients for SEP-SEQ were recruited as part of a prospective, observational study including patients presenting to an academic Emergency Department (Stanford University Hospital, Stanford, CA) with a sepsis alert. The first 1500 clinical patient samples analyzed by Karius were received sequentially by the Karius commercial laboratory up to July 31, 2018. |