Combined use of metagenomic sequencing and host response profiling for the diagnosis of suspected sepsis

Henry K. Cheng, PhD; Susanna K. Tan, MD*; Timothy E. Sweeney, MD, PhD*; Pratheepa Jeganathan, PhD*; Thomas Briese, PhD; Veda Khadka; Fiona Strouts, PhD; Simone Thair, PhD; Sudeb Dalai, MD, PhD; Matthew Hitchcock, MD; Ashrit Multani, MD; Jenny Aronson, MD; Tessa Andermann, MD; Alexander Yu, MD; Samuel Yang, MD; Susan Holmes, PhD; W. Ian Lipkin, MD; Purvesh Khatri, PhD†; David A. Relman, MD†

*These authors contributed equally to this work.

†These authors contributed equally to this work.

Affiliations: Department of Bioengineering, Stanford University, Stanford, California (H K Cheng); Division of Infectious Diseases and Geographic Medicine, Department of Medicine, Stanford University, Stanford, California (H K Cheng, S K Tan, V D Khadka, Strouts, S Dalai, Hitchcock, A Multani, J Aronson, T Andermann, A Yu, Prof Relman); Institute for Immunity, Transplantation, and Infection, and Division of Biomedical Informatics Research, Department of Medicine, Stanford University, Stanford, California (T E Sweeney, S A Thair, Prof P Khatri); Department of Statistics, Stanford University, California (P Jeganathan, Prof S Holmes); Center for Infection and Immunity and Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York (T Briese, Prof I Lipkin); Department of Emergency Medicine, Stanford University, Stanford, California (S A Thair, Prof S Yang); Department of

Microbiology & Immunology, Stanford University, Stanford, California (Relman); Infectious Diseases Section, Veterans Affairs Palo Alto Health Care System, Palo Alto, California (Prof Relman).

Corresponding Author: David A. Relman, MD, Departments of Medicine, and of Microbiology & Immunology, Stanford University School of Medicine, Encina E209, 616 Jane Stanford Way, Stanford, CA 94305 USA (relman@stanford.edu).

Word count (not including title, abstract, research in context, acknowledgment, references, tables, and figure legends)**:** 5,077

RESEARCH IN CONTEXT

Evidence before this study: Our PubMed search for articles matching the terms ("metagenomic" OR "cell-free DNA") and "infect*" in the title/abstract and using the "Human" species filter from inception to September 30, 2019 yielded 463 articles. Many proof-of-concept and validation studies illustrating how metagenomic sequencing can diagnose infections have been previously reviewed. Our search identified only nine studies which applied metagenomic shotgun sequencing to blood specimens, likely because there is a relatively low signal-to-noise ratio with this specimen type in this setting. In a study of 358 febrile sepsis patients, plasma cell-free DNA sequencing detected causative agents missed by standard-of-care testing in 15% of patients, but also detected bacterial organisms adjudicated as commensals in 10% of patients. Recently, a proof-of-concept study used machine learning to integrate metagenomic sequencing and transcriptional host response profiling to differentiate pathogens from commensal organisms in respiratory specimens, albeit with only a small derivation cohort to train host response signatures.

Added value of this study: Our 200-patient study assessed the clinical utility of combining both metagenomic sequencing and a previously-defined host response assay to diagnose sepsis. We developed a rigorous chart review approach to measure whether our assays' results could change a physician's diagnostic decision-making, without having to commit the assays into patient care. Metagenomic sequencing revealed previously-undetected and clinically relevant organisms in 17 of 200 patients, and host response profiling led at least two of three physician chart reviewers to change

their diagnostic classifications in 46 of 100 patients. We also report on potential bacterial DNA bloodstream translocation in 8 of 40 patients who were originally classified by physicians as noninfected and show how host response profiling can guide interpretation of metagenomic shotgun sequencing results. Finally, we present a statistical algorithm for contaminant removal from metagenomic sequencing data using Bayesian inference.

Implications of all the available evidence: Current diagnostic techniques are inadequate for rapid microbial diagnosis and optimal management of patients with suspected sepsis. Metagenomic sequencing, which offers the promise of hypothesis-free testing to discover new organisms that would have otherwise been missed, is already being introduced into clinical practice. However, interpretation of results from this powerful approach can be difficult, given that a large fraction of positive results represents reactivated viruses, chronic infections, commensal organisms, and contamination. Host response profiling can serve as an objective adjunct in interpreting ambiguous metagenomic sequencing results. As host response assays are introduced into clinical practice, we suggest that all patients undergoing metagenomic sequencing be simultaneously tested with one of these assays. For now, we urge clinicians to carefully interpret metagenomic sequencing results with the utmost regard for patient safety and antimicrobial stewardship.

SUMMARY

Background: Current diagnostic techniques are inadequate for rapid microbial diagnosis and optimal management of patients with suspected sepsis. We assessed the impact of metagenomic sequencing and host response profiling individually and in combination on microbiological diagnosis in these patients.

Methods: In this cohort study of 200 consecutive patients with suspected sepsis we evaluated three molecular diagnostic methods with blood specimens: 1) direct bacterial DNA detection and characterization with metagenomic shotgun next generation sequencing and contaminant sequence removal using Bayesian inference; 2) direct viral DNA and RNA enrichment and detection with viral capture sequencing; and 3) transcript-based host response profiling with a previously-defined 18-gene qRT-PCR assay. We then evaluated changes in diagnostic decision-making among three expert physicians in a chart review by unblinding our three molecular test results in a staged fashion.

Findings: Metagenomic shotgun sequencing confirmed positive blood culture results in 14 of 26 patients. In 17 of 200 patients, metagenomic sequencing and viral capture sequencing revealed organisms that were 1) not detected by conventional hospital tests within 5 days after presentation, and 2) classified as of probable clinical relevance by physician consensus. Host response profiling led at least two of three physicians to change their diagnostic decisions in 46 of 100 patients. Finally, we report on potential bacterial DNA translocation in 8 patients who were originally classified by physicians as

noninfected and show how host response profiling can guide interpretation of metagenomic shotgun sequencing results.

Interpretation: The integration of host response profiling, metagenomic shotgun sequencing, and viral capture sequencing synergistically enhances the utility of each of these approaches, and may improve the diagnosis of infections in patients with suspected sepsis.

Funding: National Institutes of Health, Chan-Zuckerberg Biohub Microbiome Initiative, and the Thomas C. and Joan M. Merigan Endowment at Stanford University.

INTRODUCTION

The early recognition and diagnosis of severe infection and sepsis is a significant clinical priority. Despite advances in microbial detection methods, clinicians typically rely on presumptive clinical diagnoses and empiric therapy with broad-spectrum antimicrobials, increasing the risks for adverse drug effects[1] and the development of antimicrobial resistance. Two emerging approaches, metagenomic sequencing and host response profiling, may each promote the rapid diagnosis of sepsis. Their use in a prospective fashion, and especially in combination, has not been adequately assessed and deserves careful study.

In theory, metagenomic sequencing can identify any microorganism to the species- or strain-level without the need for a prior hypothesis or reliance on cultivation, as long as there are nucleic acids of sufficient abundance and length from the organism(s) in the specimen. Case reports, validation, and interventional studies have highlighted the potential power of this approach[2–7]. Some methods incorporate microbial enrichment or human depletion steps in order to improve 'signal to noise' ratios[8,9]. For example, viral capture sequencing for vertebrate viruses (VirCapSeq-VERT) is a metagenomic sequencing approach that enriches for all 207 viral taxa known to infect vertebrates (including humans) with sensitivity similar to the real-time polymerase chain reaction assays currently employed in clinical microbiology laboratories[10,11].

The mere presence of specific molecular components of an infectious agent in a patient is insufficient however to incriminate the agent as the cause of that patient's disease[12,13]. For example, the presence of bacterial nucleic acids in a specimen of blood could be explained by contamination of the specimen with skin bacteria or their

DNA during collection[14], or even normal low-level translocation of commensal bacteria or their components into the bloodstream during states of health[15]. Viral sequences may represent latent or clinically-irrelevant viruses in circulating blood cells or their nucleic acids in plasma. Contamination of specimens with microbial nucleic acids from laboratory reagents at the time of specimen processing has been shown to critically affect results in the study of low-microbial biomass samples, such as blood[16]. Finally, false-positive and -negative results may reflect bioinformatic errors[17] and faulty reference databases[3,18], or other technical errors. The failure to address these same challenges in the use of other nucleic acid-based testing approaches such as multiplex pathogen PCR panels and *C. difficile* PCR testing has led to unnecessary antimicrobial treatments, delayed diagnoses, and/or detrimental patient outcomes[19–22]. The risks of these adverse outcomes are magnified with metagenomic approaches because of their broad range and the ubiquity of microbial nucleic acids.

Assessments of host response to infection offer the possibility of revealing mechanism, inciting factors, and outcome. Although well-established in clinical practice, most traditional analytes, such as acute phase reactants, are non-specific. Host RNA transcript-based profiles can provide evidence of a clinically relevant response with specificity for all infections or broad classes of infectious agents[23–27]. Thus, these methods offer complementary benefits to methods that only detect microbial signals[28]. RNA signatures that identify whether a patient is infected and the general type of infection[25,29,30] (e.g. bacterial or viral) may be able to provide results in a turnaround time that would allow for initial treatment guidance, since relevant host mRNAs are highly abundant and require relatively little sample preparation. These assays, however,

are limited as they generally do not provide species-level information about the causative agent.

We hypothesized that metagenomic sequencing and host response profiling could reveal clinically useful information that current, routine diagnostic tests fail to provide about the potential cause of suspected sepsis, and that their use in combination could prove complementary. Langelier et al. provided the first integration of these two approaches to diagnose lower respiratory tract infections[31]. In our study, we prospectively enrolled 200 consecutive adult patients who presented to the Emergency Department with suspected sepsis, as defined by a prior sepsis definition[32]. Next, we applied three molecular approaches with specimens from these 200 adult suspected sepsis patients: 1) metagenomic shotgun next-generation sequencing (mNGS) for bacteria detection in plasma specimens; 2) VirCapSeq-VERT for DNA and RNA virus detection in plasma specimens; and 3) a previously-defined human response-based transcript signature, Integrated Antibiotics Decision Module (IADM)[25], to classify bacterial infection, viral infection, and noninfection-based inflammation in whole blood samples. In addition, we developed an open-source gamma-Poisson mixture model-based Bayesian method for distinguishing blood-associated sequences (signal) from reagent-associated, contaminant DNA sequences (noise) in mNGS data. Three physicians with specialty training in infectious diseases performed chart reviews on all patients in a blinded manner and then were provided results from the three diagnostic methods in a staged fashion. We report on the added value of these methods alone and together in generating clinically relevant diagnoses.

METHODS

Subject Enrollment

This study was approved by the Stanford University Administrative Panel on Human Subjects Research (Protocols 32851, 29803, and 29733).

The patient cohort was a prospective, consecutive convenience sample of 200 patients with suspected sepsis. Plasma and PAXgene™ RNA whole blood samples were prospectively collected from adult patients presenting to the Stanford University Hospital Emergency Department (ED) who satisfied all of the following inclusion criteria:

1) Not pregnant;
2) Met 2 of 4 SIRS criteria, as defined by Bone et al.[32]; and
3) Had suspicion of infection in ED, as determined by triage nurses or other clinicians

Blood samples for this study were collected at the time of venipuncture for standard-of-care bacterial cultures during presentation under a waiver of informed consent granted by the IRB. From this patient sample bank, we then identified 200 consecutive suspected sepsis patients (spanning a 128-day period in 2016) (see Supplementary Attachment 1) who met the following additional criteria:

1) 2.5 mL of whole blood in a PAXgene RNA tube (collected as part of this study protocol) and at least 200 µL of plasma, were available;

2) blood samples underwent nucleic acid extractions without errors; and

3) no access restrictions for their electronic medical record

We note that our patient sample banking operations began before the new sepsis-3 definition[33] was released in 2016 and relied on a prior definition[32]. Thus, our enrollment

efforts did not include patients that would have otherwise been identified under the expanded sepsis-3 definition.

In addition, we collected 2.5 ml of peripheral blood in a PAXgene RNA tube from each of 10 healthy adult volunteers in the San Francisco Bay Area to serve as controls for host response profiling. Written, informed consent was obtained from each healthy volunteer prior to sampling. Inclusion criteria for these volunteers are available in the Supplementary Methods.

mNGS

DNA was extracted from 200-400 µL of plasma with the QIAamp Circulating Nucleic Acid Kit (QIAGEN). DNA extraction was performed in batches of 24, with 3-4 negative controls per batch, consisting of molecular-grade water, to monitor environmental and reagent contamination during sample processing. Libraries were prepared with the KAPA HyperPrep Kit (Roche) at the High-Throughput Sequencing and Genotyping Unit at the University of Illinois at Urbana-Champaign, and sequenced on the HiSeq 4000 (Illumina) with 2x150 nucleotide paired-end reads.

In a pilot experiment, we sequenced DNA from a plasma sample from each of 15 patients with a positive bacterial blood culture, at a depth of 40M-60M reads/sample, and 4 negative controls at a depth of 2-6M reads/sample. We then sequenced plasma samples from the other 185 patients each to a depth of 10M-52M reads using unique dual-indexed barcodes, alongside 36 negative control samples sequenced to a depth of 3M-6M reads.

After adapter sequence removal, quality trimming, and human genome sequence subtraction, Kraken[34] was run on non-human reads with a conservative alignment threshold of 0.3. Bacterial reads classified to the species-level were further analyzed with phyloseq[35].

Exploratory analysis using PCA on rank-normalized sequence reads showed possible batch-effects which may have contributed to variation in sample sequence composition (figure S1). Two distinct clusters were visualized when samples were grouped into sets based on their extraction batches: Set 1, consisting of all samples from extraction batches 1, 2, 3, 4, 11, and 12; and Set 2, which consisted of all samples from extraction batches 5, 6, 7, 8, 9, and 10. The Pilot Set, consisting of all the pilot experiment samples, clustered with Set 1. Closer examination of taxa in the negative controls of Set 1 and Set 2 revealed distinct contamination signatures, with numerous high-abundance taxa unique to each set (figure S2). We hypothesized that differences in manufacturing lots of the nucleic acid extraction kits may have caused the variation in sample sequence composition, as Glassing et al. previously reported[36]. Further details on mNGS sample preparation, bioinformatics, and exploratory analysis are provided in supplementary methods.

To distinguish blood-associated DNA sequences from contaminant sequences in plasma samples, we developed a Bayesian statistical method that leverages data from negative control samples. We ran the contaminant removal algorithm separately on the three sets of samples: Set 1, Set 2, and the Pilot Set. We analyzed the Pilot Set separately, even though it behaved similarly to Set 1, because the two sets were extracted by separate technicians and sequenced using different barcode adapters

several months apart. A detailed description for our contaminant removal algorithm is provided in appendix S1, and an open-source R package of this method is available at https://github.com/PratheepaJ/BARBI.

VirCapSeq-VERT

Nucleic acid was extracted from 150 µl of plasma using the NUCLISENS easyMAG system (bioMerieux). Sequencing libraries were prepared with the KAPA HyperPrep kit (Roche) following reverse transcription (Superscript III; Thermo Fisher) and second-strand synthesis (Klenow polymerase; New England Biolabs). Libraries were labeled with custom unique dual barcode adapters (Integrated DNA Technologies) and pooled for VirCapSeq-VERT capture hybridization (25-32 samples per pool). The VirCapSeq-VERT enriched library pools were sequenced on a HiSeq 2500 sequence analyzer (Illumina), generating 1 x 100 nucleotide single end reads.

Sequence reads were demultiplexed, Q30-filtered and assessed by RRINSEQ (v0.20.2) prior to host sequence subtraction and *de novo* assembly (MIRA v4.0). Contigs and unique singletons were subjected to homology search using MegaBLAST and the GenBank nucleotide database. Sequences not assigned at the nucleotide level were screened by BLASTx to detect divergent or potentially new viruses. Based on the BLAST results, contigs and singletons were assigned at family, genus, species and GenBank accession number level to identify the most closely related GenBank entries.

Each sample pool included a negative control consisting of Salmon nucleic acid that was processed alongside the human plasma samples. Raw read counts were normalized (reads per 10,000 host subtracted total reads) and a positive score assigned

to specimens with a result >0.2 and for which these reads did not represent a read pile-up in one position but were distributed to three or more genome regions. We did not report on viruses of the family *Anelloviridae* and GB viruses as they have no established clinical significance in humans[37,38]. Additional details on VirCapSeq-VERT sequencing and bioinformatics processing are available in supplementary methods.

Host RNA transcript profiling

We tested samples from 193 patients and 10 healthy adult volunteers with a previously described 18-gene host-response assay consisting of 1) an 11-gene set to distinguish noninfection- and infection-associated SIRS, the Sepsis MetaScore (SMS)[24]; and 2) a 7-gene set to distinguish bacterial and viral infections, the 'bacterial-viral metascore' (BVS)[25]. The Stanford Functional Genomics Facility extracted RNA from PAXgene RNA tubes using the QIAcube system (Qiagen) according to the manufacturer's recommendations, and then performed qRT-PCR for specific human transcripts in triplicate using commercial TaqMan assays on the Biomark HD platform (Fluidigm). Samples from seven patients were not profiled because of failure of PCR amplification. SMS and bacterial-viral scores were calculated as previously described[25].

Since this was the first use of qRT-PCR to measure target mRNAs, we needed to re-establish SMS and BVS cutoffs for the data generated in this study. First, physicians with subspecialty training in infectious diseases (not the three physicians of the main chart review) conducted a 'host response calibration chart review' to establish baseline classifications of infection status and type for the 193 patients with host response results. Each patient's medical records were reviewed by two physicians who were

blinded to mNGS, VirCapSeq-VERT, and host response profiling results. SMS and BVS score cutoffs were then re-established using the results from the 'derivation cohort' of 93 patients who were adjudicated as noninfected, or as having a bacterial or viral infection by physicians with evidence from standard-of-care microbiological tests. Cutoffs were set to incorporate 95% sensitivity for bacterial infections. With these score cutoffs, host response classifications of 'bacterial,' 'viral,' or 'noninfected' were generated for all 193 patients. In the main chart review, physicians were presented with plots of host response results incorporating score cutoffs for only the 100 patients in the 'test cohort'. The host response calibration chart review questions and results are available in supplementary methods and supplementary attachment 1, respectively.

Main Chart Review

We recruited three physicians with subspecialty training in infectious diseases, to perform a retrospective chart review on the 200 patients in a blinded manner. They were asked to make classifications on infection status and clinical relevance of mNGS and VirCapSeq-VERT results, in a staged fashion: 1) with only medical charts; 2) with the addition of mNGS and VirCapSeq-VERT results; and 3) with the further addition of host response results. The results of chart review are summarized in figure 1, and details are provided in supplementary methods, appendix S2, and appendix S3.

RESULTS

Patient population

We recruited 200 consecutive patients in the ED with suspected sepsis; applied mNGS, VirCapSeq-VERT, and host response profiling on blood specimens of each patient; and evaluated patient clinical records in two separate physician chart reviews, as depicted in figure 1.

Patient demographics are listed in table 1. The clinical syndromes at presentation were diverse and included fever without localizing findings (32% of patients), as well as syndromes involving the respiratory (21.5%) and genitourinary (9.5%) tracts, and intra-abdominal sites (16.5%). While these patients were enrolled because they met SIRS criteria and were suspected at the time of presentation by triage nurses in the ED of having sepsis, physicians classified 16 patients (8%) as not infected during the main chart review while blinded to mNGS, VirCapSeq-VERT, and host response profiling results. The remaining patients were classified as having bacterial infections (69 patients, 34.5%), viral infections (11 patients, 5.5%), fungal infection or coinfection (4 patients, 1%), or probable or unsure status (100 patients, 50%) (figure 2). Changes in physician classifications made after considering mNGS, VirCapSeq-VERT, and host response results are also summarized in figure 2.

Comparison of mNGS and VirCapSeq-VERT with standard-of-care microbiology

To distinguish signal from noise and remove contaminant sequences from plasma sequence data, we developed a gamma-Poisson mixture model-based Bayesian inference method. Using the 40 negative control samples, this method identified the vast majority of taxa in our dataset as contaminants (figure S3). Subsequent analyses of mNGS output were performed on contaminant-filtered data.

Bacterial sequences were identified by mNGS in plasma matching those of the species cultivated from blood collected at the same time, from the same subject in 14 of 26 patients with positive blood cultures (table 2). Interestingly, mNGS results were also concordant with the positive results of urine or sputum cultures performed within 1 day of presentation in 3 of 24 patients with negative blood cultures (table 2). To test whether sequencing depth might explain low sensitivity, we selected plasma samples from 7 patients with positive blood, urine, wound, or bronchoalveolar lavage cultures but negative mNGS results and acquired an additional 65-262M reads per sample. With these additional data, 2-111 additional sequencing reads matching the species of the isolated organism(s) were recovered in 6 of the 7 samples (table S1).

VirCapSeq-VERT high-throughput sequencing was performed on 199 of the 200 available plasma samples. One of the 200 samples failed to yield sufficient nucleic acid for analysis despite repeated extraction attempts. An average of 12 million raw reads were obtained for each of the 199 samples using this approach. In comparison to standard-of-care PCR testing on plasma, VirCapSeq-VERT confirmed the presence of cytomegalovirus DNA in 2 of 2 subjects (table 2). VirCapSeq-VERT analysis of plasma did not provide evidence of viruses that were subsequently identified by PCR tests on respiratory and stool samples in 8 patients as well as a heterophile antibody (Monospot) Epstein-Barr Virus test on 1 patient performed within 1 day of presentation (table 2, supplementary attachment 1). However, respiratory or stool samples were not tested using VirCapSeq-VERT and there were no independent molecular or culture data indicative of viremia.

To address whether organisms detected by mNGS or VirCapSeq-VERT were likely etiologic agents for the clinical presentation, our three expert physicians independently evaluated the mNGS and VirCapSeq-VERT findings in the main chart review. Of the 40 patients with organisms detected by mNGS in plasma from the day of presentation that were not identified with standard-of-care microbiological testing performed within 5 days of presentation, organisms in 14 patients were classified by physician consensus as either 'probably clinically relevant' or 'clinically relevant' (figure 3 and table 3). The addition of mNGS results led physicians to change their classification for the presence of bacterial infection in just six of these 14 patients by consensus (figure 3). The remaining eight patients were already established as known bacterial infection patients by physician consensus before mNGS results were revealed (figure 3).

The potential bacterial etiologies found only by mNGS included *Streptococcus mitis* (Pt_154), *Borrelia hermsii* (Pt_083), *Leptospira interrogans* (Pt_163), and *Haemophilus influenzae* (Pt_194) (table 3). In five patients with positive blood cultures (Pt_020, Pt_037, Pt_092, Pt_137, and Pt_145), mNGS uncovered additional organisms that were not found in the blood cultures (table 3). mNGS results and clinical details for all 40 patients with mNGS organisms not detected by standard-of-care microbiological testing are presented in table S2. We note that in 16 of these 40 patients mNGS revealed only organism(s) that represented possible contaminants (and/or misalignment errors). These organisms remained after contaminant removal because they were not known to be typically associated with humans or had relatively few reads recovered (table S2).

Of the 27 patients with viruses detected by VirCapSeq-VERT in plasma from the day of presentation that were not identified with standard-of-care microbiological tests, three patients had viruses which were classified by physician consensus as either 'probably clinically relevant' or 'clinically relevant' to the patient's presentation (figure 3 and table S3, with clinical details for each patient presented in table S4). Two of these patients had Coxsackievirus sequences, and one had probable Epstein-Barr Virus infection. VirCapSeq-VERT results from all three patients led to a change in the physicians' consensus classification for the presence of viral infection. Although they were not classified as the etiologies of the patient presentation, VirCapSeq-VERT demonstrated utility in detecting potential chronic viral infection or viral reactivation (supported by prior documented lab findings or signs and symptoms) in 18 patients with human herpesvirus 6, hepatitis C virus, hepatitis B virus, BK virus, Epstein-Barr virus, or Trichodysplasia spinulosa-associated polyomavirus. The remaining virus sequences had uncertain clinical relevance (tables S3 and S4).

Details on four patients for whom physician classifications of infection status and type were most altered by the results of mNGS and VirCapSeq-VERT are presented in table 4. These patients were found to be infected with coxsackieviruses, *Borrelia hermsii*, and *Leptospira interrogans*, which all defied clinical suspicion by the treating physicians. The patient with the greatest change in physician classification was Pt_163, a febrile male with headache and diarrhea upon return from travel to South Asia who was believed to have had a viral infection but was found to have *Leptospira interrogans* sequences in plasma by mNGS.

Impact of Host Response Profiling Results on Physician Classifications

To evaluate the impact of host mRNA response signatures on physician classifications of patients, we applied the previously-established Integrated Antibiotics Decision Module (IADM)[25] on 193 patient samples. The IADM incorporates the Sepsis MetaScore (SMS) which distinguishes noninfection- and infection-associated SIRS, and the Bacterial/Viral metaScore (BVS) which distinguishes bacterial and viral infections (figure 4A). Since this study was the first use of qRT-PCR to measure target mRNAs of the SMS and BVS, new cutoffs were set using the results from the 'derivation cohort' of 93 patients adjudicated by physicians in a separate chart review as either non-infected, or having bacterial or viral infection (figure 4B). With these new cutoffs, host response classifications of 'bacterial,' 'viral,' or 'noninfected' for all 193 patients were generated and compared against physician adjudications (figure 4C-D).

Among derivation cohort patients, bacterial, viral, and noninfection detection sensitivities were 93.8%, 45.5%, and 43.8% respectively. Bacterial, viral, and noninfection detection specificities were 70.4%, 97.5%, and 94.7%, respectively. The IADM distinguished noninfected and infected patients with an area under curve (AUC) of 0.73 (95% confidence interval [CI] 0.68-0.79), and bacterial from viral infections with an AUC of 0.89 (95% CI 0.85–0.93) (figure S4) using receiver operating characteristic (ROC) analysis.

We then examined the impact of host response profiling results on physician diagnostic decision-making of the 100 test cohort patients who were not used in setting host response score cutoffs. In 46 patients (46%), the addition of host response profiling results led at least two of three physicians to change their classification of infection

status and type (figure 5A). We also asked physicians to classify the clinical relevance of mNGS and VirCapSeq-VERT organisms first using medical charts only and then with the addition of host response results. Ten patients had at least one physician change their classification of clinical relevance of an organism revealed by mNGS or VirCapSeq-VERT upon receiving host response scores (figure 5B).

Possible bacterial DNA bloodstream translocation in patients originally classified as noninfected

In eight of the 50 patients originally classified by physician consensus as probably noninfected or noninfected, mNGS detected sequences in plasma from typical commensal organisms (table S5). Physicians noted pre-existing mucosal membrane disturbances in five of these eight patients, thus raising the possibility of bacterial DNA translocation from heavily colonized mucosal sites. For example, Pt_070, who had high abundances of sequences from more than 20 oral cavity-associated organisms in plasma, had documented gingivitis and hemoptysis. All eight patients improved after their ED visit, six of whom were not prescribed antibiotics. Host response results could have been useful to physicians for interpreting ambiguous mNGS results from these patients. However, most of these patients were in the ‘derivation cohort’ used for setting host response cutoffs, and thus did not have their host response results assessed in the main chart review (figure 4B). Nonetheless, host response profiling predicted that five of the eight patients were not infected (table S5).

Data from mNGS, VirCapSeq-VERT, host response, and physician chart reviews for all 200 patients are provided in supplementary attachment 1.

Discussion

Diagnosing infections in patients with suspected sepsis is challenging, particularly in those with multiple co-morbidities. We applied two broad-range sequencing approaches, mNGS and VirCapSeq-VERT, as well as host response profiling to a prospectively-sampled cohort of 200 adults with suspected sepsis who were enrolled in an Emergency Department. The consecutive convenience sample reflects real-world patient heterogeneity in a tertiary care hospital. We evaluated diagnostic decision-making by three infectious disease physicians as they received information from the electronic medical record, the two sequencing-based methods, and host response profiling in a staged fashion. Our results show that the sequencing methods can detect clinically relevant organisms that were missed by routine microbiological diagnostic methods, as well as other organisms that were not deemed clinically relevant. In addition, we demonstrated the potential for host response profiling to influence diagnostic decision-making and help interpret metagenomic sequencing results.

One of the most important features of unbiased, ‘shotgun’ metagenomic sequencing is that it is hypothesis-free, allowing simultaneous detection of thousands of organisms, including those difficult-to-culture. Seventeen of the 200 patients had clinically relevant organisms detected by mNGS and VirCapSeq-VERT that were not detected by standard-of-care microbiology within five days after presentation. Results from nine of these 17 patients led physicians to change their classifications of infection

status and type. For example, patient Pt_083 presented with fever after travel to the Sierra mountains and was presumed to have a urinary tract infection by treating physicians. However, this patient was determined by mNGS to have tick-borne relapsing fever due to *Borrelia hermsii*. Our positivity rate was comparable to other clinical metagenomics studies as reviewed by others[39,40]. For example, in a study of 204 meningitis and encephalitis patients diagnoses in 13 of them were made solely by metagenomic sequencing with CSF samples, with an impact on patient management in 7 of the 13[41]. In a study of cell-free plasma in 358 febrile sepsis patients, 15% of patients had probable causal pathogens detected solely by metagenomic sequencing[5]. It should be noted that metagenomic sequencing can provide a 53-hour turnaround time[5], which is shorter than standard-of-care tests in some situations.

Contaminant sequence identification and computational removal represents one of the greatest barriers to expanding the clinical application of metagenomic sequencing, especially in specimens with low microbial biomass such as blood. The gamma-Poisson mixture model-based Bayesian inference approach that we have introduced here offers an important advance in addressing this challenge. In our implementation, we assumed that DNA sequences in plasma included those of contaminants. We then inferred the true 'intensity' of DNA sequences in a plasma sample that might be attributed to 'true' blood-associated nucleic acids. This method adds to others available to researchers for contaminant removal[4,36,37]. For example, for studies with fewer than three negative control samples, simply subtracting species based on their presence or abundance in negative controls[42] may be most appropriate. While the decontam[43] method is not well-suited for our data because it assumes that

samples have a relatively higher biomass than controls, and that each taxon is either a contaminant or 'true' but not both, it can be very helpful for 16S rRNA gene amplicon sequence data from other kinds of samples.

As metagenomic sequencing enters clinical practice, it is important to recognize the potential of this powerful approach to reveal true signals, as well as clinically-irrelevant sequences which can occur because of translocation of microbial nucleic acids from heavily colonized body sites, reactivation of latent viruses, or contamination of laboratory reagents or specimen collection devices. Virus sequences from 18 of 27 patients with positive VirCapSeq-VERT results were associated with chronic infections or viral reactivations that were not clinically relevant to the patient's presentation. Additionally, eight of 50 patients who were originally classified as noninfected or probably noninfected had bacterial organisms detected by plasma mNGS. Five of these eight patients improved without antibiotics. Clinicians are accustomed to the importance of clinical-pathological correlations for establishing the relevance of laboratory findings. With the advent of sensitive molecular diagnostic technologies, this challenge will only grow. Indeed, Blauwkamp et al. detected organisms adjudicated as 'commensal' in 36 of 358 febrile sepsis patients (10.0%) and as 'viral reactivation' in 10 of 358 patients (2.8%) with metagenomic sequencing of cell-free plasma DNA[5]. They suggested that sequence abundance and overall clinical picture should be considered while assessing clinical relevance of metagenomics results.

Our data illustrate the utility of transcriptional host response signatures, as an objective adjunct in guiding the interpretation of mNGS results and avoiding misdiagnosis and unnecessary treatment. Our results add to those of Langelier et al.,

who combined host response and metagenomic sequencing to diagnose lower respiratory tract infections[31] using a different approach from ours. Their study included a sophisticated machine-learning-based integration of the complementary approaches, using a training set of 20 patients to generate signatures which identified infectious etiologies vs. commensal organisms in respiratory metagenomic sequencing results. In our study, host response signatures for identifying bacterial vs. viral vs. noninfected patients were previously trained on datasets from over 2,000 patients[25] representing a wide diversity of infectious etiologies. Additionally, we focused on measuring the clinical utility of having physician chart reviewers integrate metagenomic sequencing and host response profiling results into their clinical decision-making.

The major limitation of our mNGS protocol was suboptimal sensitivity, which was explained in part by the choice of sequencing depth. Low numbers of patients with known systemic viral infections limited our ability to assess the performance characteristics of VirCapSeq-VERT. Bioinformatic errors[17] from contaminant, misannotated, or missing genomes in microbial databases, and other technical limitations such as index-hopping[44], may have led to false-negative and false-positive findings. Recent efforts to curate databases[45], manufacture contaminant-free extraction kits, and enrich for microbial sequences[6,8] are steps in the right direction to prepare metagenomic sequencing for routine clinical use, but much work remains to be done.

Our bacterial contaminant sequence identification method did not subtract all contaminant sequences in our mNGS dataset. Increasing the number of negative control samples in every extraction batch could aid in profiling the large diversity of contaminating taxa and thus enhance contaminant sequence removal. Furthermore, our

negative control samples were only suited for identifying extraction reagent contaminants. We were not positioned to account for skin-associated contaminants or spurious sample-to-sample cross-contaminants.

The host response profiling assay classified many viral and noninfected patients as bacterial. One possible reason for these misclassifications was the strict dichotomous cutoffs that we used to distinguish infected vs. noninfected cases, and viral vs. bacterial infections. Reporting results with numeric values rather than dichotomous cutoffs will allow better weighting of these scores in patient assessments. Another reason for the misclassifications was the need to re-establish host response score cutoffs for this study's qRT-PCR platform and the small number of known viral patients with which to do so. Further work is needed to establish and lock cutoffs, validate on additional patient populations, and quantify test characteristics.

We believe that host response profiling and shotgun sequencing will soon achieve turnaround times of less than 90 minutes and 24 hours, respectively. In our chart review, physicians had access to the patients' full medical chart histories, including test results that only became available several days after presentation. If it were possible to limit the review to the first 90 minutes of each case history, we expect that there would have been greater changes in diagnostic decision-making across the stages of our chart review.

A central limitation in evaluating new diagnostic tools is the lack of a gold standard. We did our best to address this using expert physicians in a staged chart review. Fundamentally, it is impossible to determine whether the changes in patient classification were correct. However, the measurement of a diagnostic tool's ability to

change clinical decision-making, rather than just a comparison of its results to standard-of-care testing, is a valuable component of establishing clinical utility. An important secondary finding was that clinicians had varying levels of trust in these new diagnostic tools. In conclusion, our proof-of-concept study on a consecutive, prospectively-sampled patient cohort suggests that integrating host response profiling with metagenomic sequencing may synergistically enhance the utility of each assay, and ultimately, the diagnosis of patients with suspected sepsis.

Acknowledgements

We thank Ian Brown, Patrice Callagy, Cheryl Bucsit, and Adele Araya of the Stanford Emergency Department for facilitating the collection of samples for this project. We thank members of the Relman Lab, and in particular, Stephen J. Popper, Eitan Yaffe, Christine L. Sun, Daniela S. Goltsman, and Natalie Campen for valuable advice, feedback, and general assistance. We also thank members of the Lipkin Lab, Joel A. Garcia, Nishit P. Bhuva, and Lokendrasingh Chauhan for technical assistance, and Bohyun Lee and Komal Jain for bioinformatics support, Kelly Murphy (Stanford Emergency Department) for advice and assistance, and John Coller and Xuhuai Ji at the Stanford Functional Genomics Facility for their technical assistance with the Fluidigm assay. We thank Alvaro Hernandez and Chris Wright at the High-Throughput Sequencing and Genotyping Unit at the University of Illinois at Urbana-Champaign for their assistance and support with DNA library preparation and sequencing. This work was supported by NIH U19 AI109761 (D.A.R., W.I.L.), Chan-Zuckerberg Biohub

Microbiome Initiative (D.A.R.), and the Thomas C. and Joan M. Merigan Endowment at Stanford University (D.A.R.).

Conflict of Interest Disclosures: Dr. Cheng, Dr. Thair, and Dr. Sweeney are employees of Inflammatix. Dr. Strouts is an employee of Cepheid. Dr Dalai is an employee of Karius. Dr. Lipkin is an advisor to Pathogenica. Dr Khatri is an advisor to Inflammatix. Dr Relman is an advisor to Arc Bio and Karius.

Table 1. Clinical characteristics of patient population.

	Patients (%)
Infection Status and Type* (n=200)	
Noninfected	16 (8%)
Bacterial Infection	69 (34.5%)
Viral Infection	11 (5.5%)
Fungal Infection	1 (0.5%)
Bacterial-Viral Coinfection	2 (1%)
Bacterial-Fungal Coinfection	1 (0.5%)
Probable or Unsure	100 (50%)
Type of Syndrome†	
Systemic‡	64 (32%)
Respiratory	43 (21.5%)
Genitourinary	39 (19.5%)
Intra-Abdominal	33 (16.5%)
Skin	10 (5%)
Ear, Nose & Throat	5 (2.5%)
Bone/Joint	3 (1.5%)
Central Nervous System	3 (1.5%)
Immune Status	
Immunocompromised, due to Cancer/Chemotherapy	68 (34%)
Immunocompromised, due to Transplant	15 (7.5%)
Immunocompromised, due to Immunosuppressing Drugs	14 (7%)
Immunocompromised, due to Cancer/Chemo and Transplant	1 (0.5%)
Immunocompromised, due to Congenital Disorder	1 (0.5%)
Immunocompetent	101 (50.5%)
Neutropenia	
Neutropenic	18 (9%)
Not Neutropenic	182 (91%)
Admission Status	
Admitted to ICU	13 (6.5%)
Admitted to Floor	134 (67%)
Sent Home from ED	53 (26.5%)
Sex	
Female	96 (48%)
Male	104 (52%)
Age (years)	
Median (Interquartile Range)	51.5 (35-68)

*Information on infection status and type was collected from the main chart review while physicians were blinded to mNGS, VirCapSeq-VERT, and host response results. The 'probable or unsure' category includes all patients without a definite known diagnosis, including those classified as probable noninfected, probable bacterial, probable viral, and unsure. Classifications had consensus agreement by at least two of three physicians. Details on how each patient was placed into each category is provided in the Supplementary Methods. All other information in this table was extracted by a single physician at the completion of our study.

†Refers to the localization of signs and symptoms of patients at presentation.

‡Systemic refers to non-localized infection, sepsis, 'viral syndrome', fever with neutropenia, post-operative fever, and/or SIRS findings on presentation not related to infection, such as those associated with malignancies (e.g., leukemia, lymphoma, and metastatic tumors), and autoimmune disorders.

Table 2. Performance of plasma mNGS and VirCapSeq-VERT versus clinically-indicated standard-of-care microbiology

		Standard of Care Microbiology*			
		Positive Bacterial Tests		Positive Viral Tests	
		Blood Culture (n=25)	Other Body Site Cx, PCR or immunoassay (with Negative Blood Culture)† (n=24)	Plasma PCR‡ (n=2)	Respiratory or Stool PCR, Monospot Antibody Test§ (n=9)
Plasma mNGS (Bacteria) or VirCapSeq-VERT (Viruses)	Confirmed Infectious Agent(s)	14	3	2	0
	No Confirmed Infectious Agent	11	21	0	9

*Includes only tests performed within 1 day of presentation and determined to be clinically relevant to the patient's presentation by physician chart review. All blood cultures were performed at the same time as the blood draw for our study.
†Tests include the following: aerobic and/or anaerobic bacterial cultures of sputum, bronchoalveolar lavage fluid, urine, wound, and intra-abdominal abscess, and perianal abscesses; stool PCR for *Salmonella enterica* and upper throat swab PCR for *Streptococcus dysgalactiae* ssp*. Equisimilis*; and upper throat swab rapid enzyme immunoassay test for group A beta hemolytic *Streptococcus*
‡Includes plasma PCRs for CMV.
§Includes upper nasopharyngeal swab PCRs for influenza, coronavirus, respiratory syncytial virus, rhinovirus; stool PCR for norovirus; and Epstein-Barr virus Heterophile Antibody (Monospot) Test

Table 3. Patients with 'clinically relevant' or 'probably clinically relevant' mNGS and VirCapSeq-VERT organism(s) not detected by hospital tests

Patient*	mNGS or VirCapSeq-VERT Organism†	Standard-of-Care Microbiology Within 5 Days After Presentation	Host Response	Final Clinical Diagnosis
Pt_003	*Coxsackievirus A6 (33,309,209 / 8,185.44)*	All negative.	Bacterial	Viral Syndrome
Pt_006	*Escherichia coli* (37, 26, 12)	All negative.	Bacterial	Bacteremia – Catheter-associated
Pt_020	*Klebsiella pneumoniae* (85, 69, 4), *Escherichia coli* (45, 33, 12)	Blood Cx (D1): *Escherichia coli*; Abdominal Wound Cx (D1): *Clostridium striatum.*	PCR Error	Bacteremia – Source, Unclear
Pt_025	*Coxsackievirus B5 (30,324 / 181.31)*	All negative.	Bacterial	Viral Syndrome
Pt_037	*Prevotella denticola* (104, 86, 2), *Porphyromonas asaccharolytica* (73, 58, 2), *Fusobacterium nucleatum* (61, 47, 3)	Blood Cx (D1): *Escherichia coli*, *Streptococcus anginosus* group; Perianal Abscess Cx (D1): *Streptococcus anginosus* group; …‡	Bacterial (Derivation Cohort)¶	Bacteremia – Source, Abscess
Pt_041	*Fusobacterium nucleatum* (22, 14, 5)	All negative§.	Bacterial	Intra-Abdominal Abscess
Pt_057	*Escherichia coli (*119, 100, 5)	Blood Cx (D1): CoNS	Noninfected (Derivation Cohort)¶	Allograft Rejection
Pt_076	*Epstein-Barr Virus (11,422 / 140.30)*	All negative.	Bacterial	UTI vs. Malignancy – Lymphoma
Pt_083	*Borrelia hermsii* (306, 273, 0)	Urine Cx (D1): CoNS	Bacterial	Tick Borne Relapsing Fever
Pt_092	*Streptococcus agalactiae* (21, 13, 0), *Streptococcus anginosus* (9, 4, 2)	Blood Cx (D1): *Streptococcus agalactiae* (Group B), *Escherichia coli*; Urine Cx (D1): *Lactobacillus* spp., *Escherichia coli*	Bacterial (Derivation Cohort)‖[e]	Bacteremia - Source Skin
Pt_113	*Klebsiella pneumoniae* (240, 213, 4)	All negative.	Bacterial (Derivation Cohort)¶	Ulcerative Colitis Flair
Pt_126	*Morganella morganii* (29, 20, 2)	All negative.	Bacterial (Derivation Cohort)¶	Bacteremia - Source Prostate
Pt_137	*Escherichia coli* (207, 181, 12), *Clostridium perfringens* (101, 83, 2)	Blood Cx (D1): *Escherichia coli*, *Klebsiella oxytoca*	Bacterial (Derivation Cohort)¶	Bacteremia - Source Intra-Abdominal
Pt_145	*Enterobacter hormaechei* (123, 103, 2), *Klebsiella pneumoniae* (26, 17, 4), *Enterobacter cloacae* (16, 9, 2)	Blood Cx (D1): *E. cloacae* complex, *Streptococcus anginosus* group.	Bacterial (Derivation Cohort)¶	Bacteremia - Source Intra-Abdominal
Pt_154	*Streptococcus mitis* (19, 12, 6)	Blood Cx (D1): CoNS	PCR Error	Febrile Neutropenia - Unclear Etiology
Pt_163	*Leptospira interrogans* (214, 186, 0)	All negative.	Bacterial	Leptospirosis
Pt_194	*Haemophilus influenzae* (35, 24, 8)	All negative.	Bacterial	Pneumonia vs. Radiation Pneumonitis

*Only organisms that were classified as 'clinically relevant' or 'probably clinically relevant' to the patient's presentation by consensus in the main chart review were included in this Table.

†For bacterial organisms, numbers in parentheses represent raw reads, estimated lower limit for the intensity of blood-associated reads, and estimated upper limit for the intensity of contaminant reads in controls. For viruses, numbers in parentheses represent (Raw Reads / Reads per 10,000 Host Subtracted Reads).
‡Additional results from this patient include: Blood Cx (D2): *Streptococcus anginosus* group; Perianal Abscess Fluid Cx (D3): *Bacteroides fragilis* group; Perianal Abscess Fluid Cx (D3): *Escherichia coli*, *Streptococcus anginosus* group
§Note: The following test was also identified: Intra-abdominal fluid Cx (D34):*Fusobacterium nucleatum*, *Prevotella* spp., *Citrobacter freundii* complex, *Enterococcus faecium*, Rare number *Klebsiella pneumoniae*.
¶Indicates patients from derivation cohort who had their host response results used to re-establish cutoffs for host response scores (see Figure 4B).
CoNS = Coagulase Negative Staphylococcus, Cx = Culture, d = days

Table 4. Clinical details on four cases in which mNGS and VirCapSeq-VERT had the greatest influence in changing physician classifications

*Patient**	*Presentation*	*SOC Micro-biology*	*Anti-biotics?*	*mNGS and VirCapSeq-VERT*	*Host Response*	*Chart Review Classifications*
Pt_003	Male with fever, myalgia, and erythematous lesions suspected from bug bites. Presumed cellulitis. Discharged home after IV fluids, no follow-up data.	Blood cx neg.	Yes	mNGS: Negative VirCapSeq-VERT: Coxsackievirus A6	Bacterial	
Pt_025	Female with fever, vomiting, diarrhea, and headache. Presumed UTI and/or viral syndrome. Discharged home after IV fluids, no follow-up data.	Blood and urine cx neg. Pyuria on urinalysis.	Yes	mNGS: Negative VirCapSeq-VERT: Coxsackievirus B5	Bacterial	
Pt_083	Male with fever, dry cough, and BPH-related urinary retention returning from field work in a mountainous region. Presumed urosepsis. Discharged home; improved on 3d follow-up.	Blood and urine cx neg. Pyuria on urinalysis.	Yes	mNGS: *Borrelia hermsii* VirCapSeq-VERT: Negative	Bacterial	
Pt_163	Male with fever, headache, and diarrhea upon return from Sri Lanka. Presumed viral infection. Discharged home under strict return precautions; symptoms continued but improved on 8d follow-up.	Blood cx, dengue, and malaria tests neg.	No	mNGS: *Leptospira interrogans* VirCapSeq-VERT: Negative	Bacterial	

	Rev. 1			Rev. 2			Rev. 3		
	PI	PII	PIII	PI	PII	PIII	PI	PII	PIII
Infected?	3	3	3	3	4	4	4	5	5
Bacterial	3	3	3	2	2	3	4	1	2
Viral?	3	3	3	3	4	4	2	5	5

	Rev. 1			Rev. 2			Rev. 3		
	PI	PII	PIII	PI	PII	PIII	PI	PII	PIII
Infected?	2	3	3	4	5	5	5	5	5
Bacterial	2	3	3	2	1	1	4	2	2
Viral?	1	3	3	4	5	5	2	5	5

	Rev. 1			Rev. 2			Rev. 3		
	PI	PII	PIII	PI	PII	PIII	PI	PII	PIII
Infected?	4	5	5	5	5	5	4	5	5
Bacterial	4	5	5	3	5	5	4	5	5
Viral?	2	1	1	3	1	1	1	1	1

	Rev. 1			Rev. 2			Rev. 3		
	PI	PII	PIII	PI	PII	PIII	PI	PII	PIII
Infected?	4	5	5	5	5	5	5	5	5
Bacterial	2	4	4	2	5	5	2	4	5
Viral?	4	2	2	4	1	1	4	2	1

Legend	**Phase I (PI)**	**Phase II (PII)**	**Phase III (PIII)**	**No**	**Probably No**	**Unsure**	**Probably Yes**	**Yes**
	Medical Charts Only	+mNGS, VirCapSeq-VERT	+Host Response	1	2	3	4	5

*Patients were identified by the following criteria: 1) An organism was revealed by mNGS or VirCapSeq-VERT that was not previously detected by standard-of-care microbiology, and 2) physicians classified the organism as clinically relevant or probably clinically relevant by consensus, and 3) mNGS or VirCapSeq-VERT results led at least one physician to increase their clinical suspicion for bacterial or viral infection by at least two points on our five-point scale to "Yes."

SOC = Standard-of-care; mNGS = metagenomic next generation sequencing; cx = culture; BPH = benign prostate hypertrophy, Rev = physician chart reviewer

Figure Legends

Figure 1. Study design. We applied three diagnostic approaches to our cohort of 200 adult patients with suspected sepsis: 1) direct bacterial DNA detection and characterization with plasma metagenomic 'shotgun' next generation sequencing (mNGS) and contaminant sequence identification using Bayesian inference; 2) direct viral DNA and RNA enrichment and detection with plasma viral capture sequencing (VirCapSeq-VERT); and 3) transcript-based host response profiling with a previously-defined 18-gene qRT-PCR assay on whole blood. Additionally, two separate chart reviews were performed. First, a 'Host Response Calibration Chart Review' established baseline diagnoses of all patients for the sole purpose of calibrating host response cutoffs. Second, a 'Main Chart Review' evaluated changes in diagnostic decision-making among three expert physicians in a chart review by unblinding our three molecular test results in a staged fashion.

Figure 2. Introduction of mNGS, VirCapSeq-VERT, and host response profiling led to changes in physician classifications. At each phase of our main chart review, patients were either assigned with high confidence to one of four diagnostic categories by a panel of three physicians or classified to have only a probable (e.g. probably bacterial, probably noninfected) or unsure diagnosis. Physicians did not evaluate host response scores from seven patients who had host response assay fail due to amplification errors, and 93 patients who had host response scores used to set cutoffs.

For Phase III, the same classification from Phase II was kept for patients who did not have host response scores for evaluation.

Figure 3. Clinical utility of positive mNGS and VirCapSeq-VERT results. Tables and Venn diagrams illustrate the number of patients with clinically relevant mNGS and VirCapSeq-VERT results which reveal the etiologies of patient presentations and change diagnostic decision-making. Patients were grouped by those who had positive (A) mNGS and (B) VirCapSeq-VERT results which 1) were not detected by standard-of-care microbiology performed within 5 days after presentation, 2) were classified as 'clinically relevant' or probably 'clinically relevant' to the patient's presentation by physician consensus while blinded to host response results during the main chart review, and 3) led physicians to have a consensus change in classification for presence of bacterial/viral infection by at least 1 point on our 5-point Likert scale. Patients were segmented into five groups using consensus classifications made during the main chart review when physicians were blinded to mNGS, VirCapSeq, and host response results. Details for all patients with positive mNGS and VirCapSeq-VERT results which were not previously detected by standard-of-care microbiological testing are presented in Table S2 and S4, respectively.

Figure 4. Host response calibration. (A) Schematic for two numeric scores of host response assay. The Sepsis MetaScore (SMS) distinguishes noninfection- and infection-associated SIRS, and the Bacterial/Viral metaScore (BVS) distinguishes bacterial and viral infections. (B) SMS and BVS score cutoffs were re-established using

the results from the 'derivation cohort' of 93 of 193 patients for whom physicians adjudicated as noninfected, bacterial, or viral in a separate 'host response calibration chart review.' With these score cutoffs, host response classifications of 'bacterial,' 'viral,' or 'noninfected' were generated for all 193 patients. In the main chart review, physicians were presented with plotted host response results with score cutoffs (C) for only the 100 patients in the 'test cohort' to interpret. (C) Distribution of scores and cutoffs for the host response assay. A higher SMS indicates a higher chance of infection over noninfection, and a higher BVS indicates a higher chance of viral infection over bacterial infection. (D) Confusion matrix for the host response assay vs. adjudicated physician classifications. The following characteristics were calculated from derivation cohort patients with total n = 93: bacterial infection sensitivity, 93.9%; bacterial infection specificity, 73.2%; viral infection sensitivity, 44.4%; viral infection specificity, 93.8%; noninfected sensitivity, 31.3%; noninfected specificity, 94.8%.

Figure 5. Clinical utility of host response results. In our main chart review, physicians reviewed host response results of the 100 patients in the validation cohort. (A) Table and Venn diagrams illustrate the number of patients for whom the introduction of host response results led physicians to change their classification(s) of infections status and type by at least one or two points in our main chart review. These classifications included whether the physicians believed the patient had 1) an infection, 2) a bacterial infection, 3) a viral infection, 4) a fungal infection, or 5) a parasitic infection. Response options were on a five-point scale (No-Probably No-Unsure-Probably Yes-Yes). (B) Twenty-five of the 100 test cohort patients had positive mNGS

or VirCapSeq-VERT results. After reviewing host response results, ten of the 25 patients had physicians change their classification of the mNGS and VirCapSeq-VERT organism's clinical relevance to the patient's presentation. Physician classifications of the ten patients' mNGS and VirCapSeq-VERT organisms before and after the introduction of host response results are presented in the bottom table.

Bibliography

1 Tamma PD, Avdic E, Li DX, Dzintars K, Cosgrove SE. Association of adverse events with antibiotic use in hospitalized patients. *JAMA Intern Med* 2017; **177**: 1308–15.

2 Grumaz S, Stevens P, Grumaz C, *et al.* Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med* 2016. DOI:10.1186/s13073-016-0326-8.

3 Simner PJ, Miller S, Carroll KC. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin Infect Dis* 2018; **66**: 778–88.

4 Hong DK, Blauwkamp TA, Kertesz M, Bercovici S, Truong C, Banaei N. Liquid biopsy for infectious diseases: sequencing of cell-free plasma to detect pathogen DNA in patients with invasive fungal disease. *Diagn Microbiol Infect Dis* 2018; : 4–7.

5 Blauwkamp TA, Thair S, Rosen MJ, *et al.* Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat Microbiol 2019* 2019; : 1.

6 Miller S, Naccache SN, Samayoa E, *et al.* Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res* 2019; : 1–12.

7 Stryke D, Briggs B, Langelier C, *et al.* Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. 2019; : 2327–40.

8 Metsky HC, Siddle KJ, Gladden-Young A, *et al.* Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. *Nat Biotechnol* 2019; **37**: 160–8.

9 Gu W, Miller S, Chiu CY. Clinical Metagenomic Sequencing for Pathogen Detection. 2019; : 317–36.

10 Williams SH, Cordey S, Bhuva N, *et al.* Investigation of the Plasma Virome from Cases of Unexplained Febrile Illness in Tanzania from 2013 to 2014: a Comparative Analysis between Unbiased and VirCapSeq-VERT High-Throughput Sequencing Approaches. *mSphere* 2018; **3**. DOI:10.1128/mSphere.00311-18.

11 Briese T, Kapoor A, Mishra N, *et al.* Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *MBio* 2015; **6**: e01491-15.

12 Lipkin WI. The changing face of pathogen discovery and surveillance. *Nat Rev Microbiol* 2013; **11**: 133–41.

13 Fredricks DN, Relman DA. Sequence-Based Identification of Microbial Pathogens: a Reconsideration of Koch's Postulates. *Clin Microbiol Rev* 1996; **9**: 18–33.

14 Hall KK, Lyman JA. Updated review of blood culture contamination. *Clin Microbiol Rev* 2006; **19**: 788–802.

15 Tomás I, Diz P, Tobías A, Scully C, Donos N. Periodontal health status and bacteraemia from daily oral activities: Systematic review/meta-analysis. *J Clin Periodontol* 2012; **39**: 213–28.

16 Salter SJ, Cox MJ, Turek EM, *et al.* Reagent and laboratory contamination can

critically impact sequence-based microbiome analyses. *BMC Biol* 2014; **12**: 87.

17 Peabody MA, Van Rossum T, Lo R, Brinkman FSL. Evaluation of shotgun metagenomics sequence classification methods using in silico and in vitro simulated communities. *BMC Bioinformatics* 2015; **16**. DOI:10.1186/s12859-015-0788-5.

18 Laurence M, Hatzis C, Brash DE. Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLoS One* 2014; **9**: 1–8.

19 Polage CR, Gyorke CE, Kennedy MA, *et al.* Overdiagnosis of clostridium difficile infection in the molecular test era. *JAMA Intern Med* 2015; **175**: 1792–801.

20 Demogines A, Fouch S, Everhart K, *et al.* Multi-Center Clinical Evaluation of a Multiplex Meningitis / Encephalitis PCR Panel for Simultaneous Detection of Bacteria , Yeast , and Viruses in Cerebrospinal Fluid Specimens. *J Clin Microbiol* 2015; **54**: 2251–61.

21 Gomez CA, Pinsky BA, Liu A, Banaei N. Delayed Diagnosis of Tuberculous Meningitis Misdiagnosed as Herpes Simplex Virus- 1 (HSV-1) Encephalitis with the FilmArray Syndromic PCR panel. *Open Forum Infect Dis* 2016; : ofw245.

22 Morgan DJ, Malani P, Diekema DJ. Diagnostic stewardship - leveraging the laboratory to improve antimicrobial use. *JAMA - J Am Med Assoc* 2017; **318**: 607–8.

23 Boldrick JC, Alizadeh AA, Diehn M, *et al.* Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc Natl Acad Sci* 2002; **99**: 972 LP – 977.

24 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. *Sci Transl Med* 2015; **7**: 1–16.

25 Sweeney TE, Wong HR, Khatri P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. *Sci Transl Med* 2016; **8**: 346ra91-346ra91.

26 Manger I, Relman D. How the host ' sees ' pathogens: global gene expression responses to infection Ian D Manger and David A Relman. *Curr Opin Immunol* 2000; **12**: 215–8.

27 Nau GJ, Richmond JFL, Schlesinger A, Jennings EG, Lander ES, Young RA. Human macrophage activation programs induced by bacterial pathogens. *Proc Natl Acad Sci* 2002; **99**: 1503–8.

28 Relman DA. New technologies, human-microbe interactions, and the search for previously unrecognized pathogens. *J Infect Dis* 2002; **186 Suppl**: S254–8.

29 Tsalik EL, Henao R, Nichols M, *et al.* Host gene expression classifiers diagnose acute respiratory illness etiology. *Sci Transl Med* 2016; **8**: 322–11.

30 Herberg JA, Kaforou M, Wright VJ, *et al.* Diagnostic test accuracy of a 2-transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. *JAMA - J Am Med Assoc* 2016; **316**: 835–45.

31 Langelier C, Kalantar KL, Moazed F, *et al.* Integrating host response and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults. *Proc Natl Acad Sci* 2018; **115**: E12353 LP-E12362.

32 Bone R, Balk R, Cerra F, *et al.* American College of Chest Physicians/Society of

Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med* 1992; **20**: 864–74.

33 Singer M, CS D, Seymour C, al et. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016; **315**: 801–10.

34 Wood DE, Salzberg SL. Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014; **15**. DOI:10.1186/gb-2014-15-3-r46.

35 McMurdie PJ, Holmes S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* 2013; **8**. DOI:10.1371/journal.pone.0061217.

36 Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog* 2016; **8**: 1–12.

37 Hino S, Miyata H. Torque teno virus (TTV): current status. *Rev Med Virol* 2007; **17**: 45–57.

38 Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P. The GB viruses: A review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. *J Gen Virol* 2011; **92**: 233–46.

39 Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet* 2019; **20**: 341–55.

40 Mitchell SL, Abmm D, Simner PJ, Abmm D. Sequencing in Clincial Microbiology: Are We There Yet? *Clin Lab Med* 2019; **39**: 405–18.

41 Wilson MR, Sample HA, Zorn KC, *et al.* Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl J Med* 2019; **380**: 2327–40.

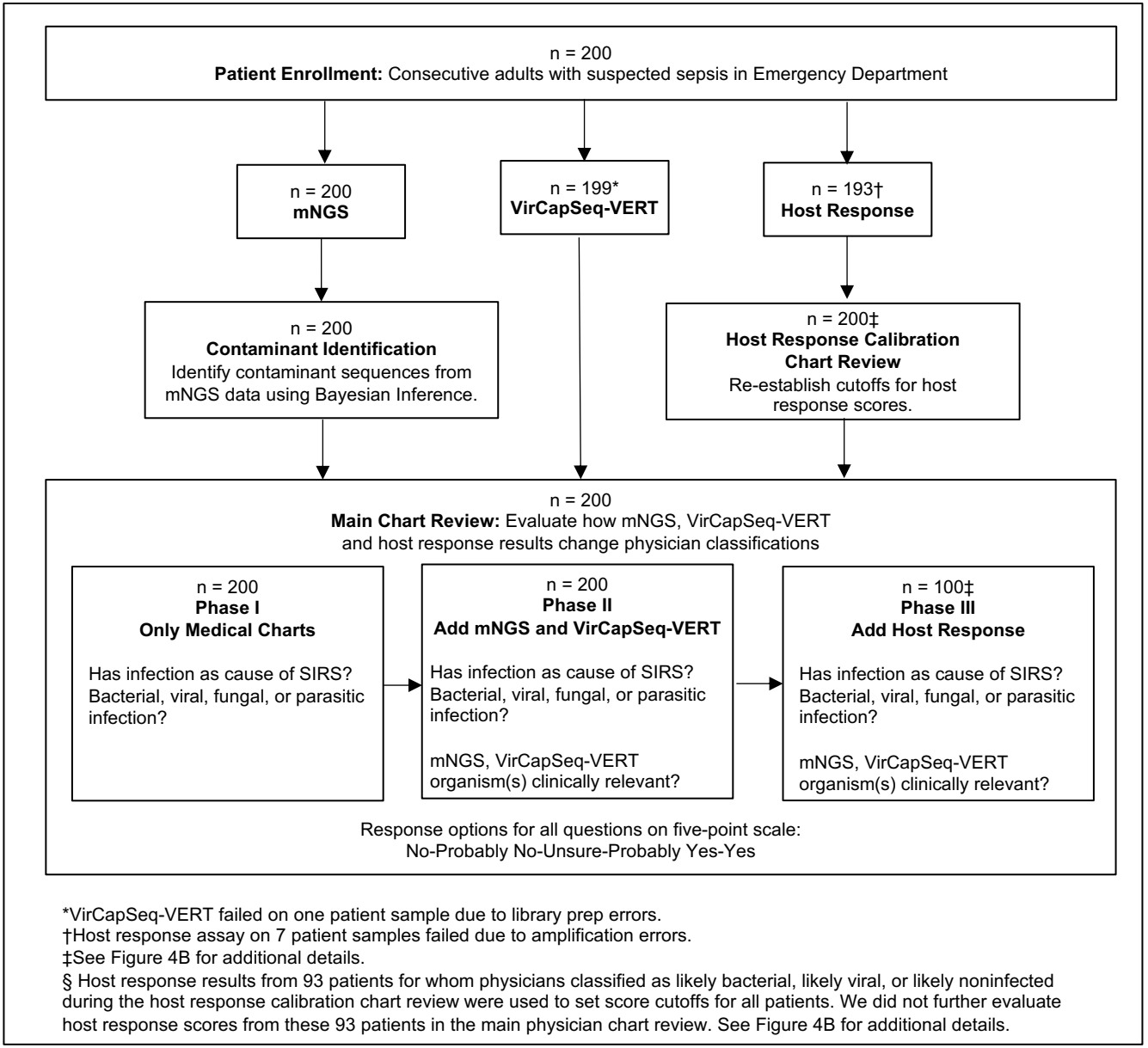
42 Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock G. Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch Pathol Lab Med* 2017; **141**: 776–86.

43 Davis NM, Proctor DiM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018; **6**: 1–14.

44 van der Valk T, Vezzi F, Ormestad M, Dalén L, Guschanski K. Index hopping on the Illumina HiseqX platform and its consequences for ancient DNA studies. *Mol Ecol Resour* 2019; **0**. DOI:10.1111/1755-0998.13009.

45 Kirstahler P, Bjerrum SS, Friis-Møller A, *et al.* Genomics-Based Identification of Microorganisms in Human Ocular Body Fluid. *Sci Rep* 2018; **8**: 4126.

Fig. 1



*VirCapSeq-VERT failed on one patient sample due to library prep errors.
†Host response assay on 7 patient samples failed due to amplification errors.
‡See Figure 4B for additional details.
§ Host response results from 93 patients for whom physicians classified as likely bacterial, likely viral, or likely noninfected during the host response calibration chart review were used to set score cutoffs for all patients. We did not further evaluate host response scores from these 93 patients in the main physician chart review. See Figure 4B for additional details.

Fig. 2

Phase I (Only Medical Charts)
Physician Classifications by Consensus (N = 200)

Noninfected	Bacterial	Viral	Fungal or Co-Infection	Probable or Unsure
16	69	11	4	100

↓ ↓

mNGS (N = 200)
53 positive, 147 negative

VirCapSeq-VERT (N = 199)
27 positive, 173 negative

↓ ↓

Phase II (+ mNGS and VirCapSeq-VERT)
Physician Classifications by Consensus (N = 200)

Noninfected	Bacterial	Viral	Fungal or Co-Infection	Probable or Unsure
12	71	11	4	102

↓

Host Response Profiling (N = 100)
13 Noninfected, 76 Bacterial, 11 Viral

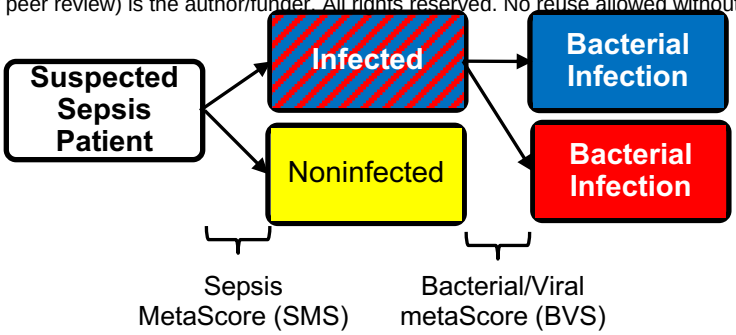
↓

Phase III (+ Host Response, if Available)
Physician Classifications by Consensus (N = 200)

Noninfected	Bacterial	Viral	Fungal or Co-Infection	Probable or Unsure
10	84	14	3	89

Fig. 4

A



B

Objective: Re-establish SMS and BVS cutoffs

Previous cutoffs were not compatible with the qRT-PCR platform employed in this study.

n = 200

Host Response Calibration Chart Review

Expert physicians* established baseline diagnoses from medical charts to calibrate host response scores.

Questions: Has infection as cause of SIRS? Has bacterial, viral, fungal, or parasitic infection confirmed by hospital tests?

* Physicians different from three physicians of main chart review.

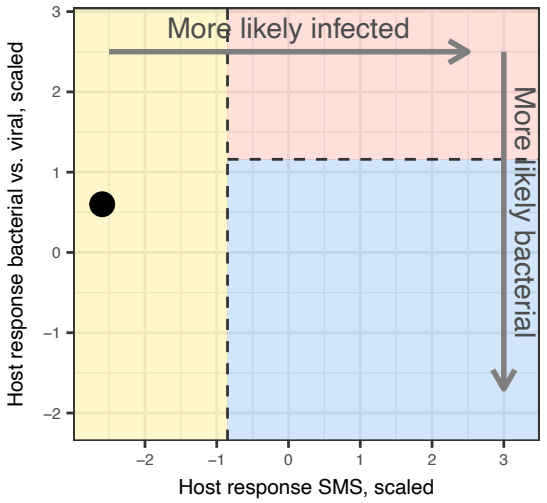
Adjudicated Physician Classification	Number Patients*
Noninfected	32
Bacterial†	52
Viral	9
Bacterial-Viral Coinfection	2
Infected, Unknown Etiology	33
Indeterminate	65

Derivation Cohort (n=93) Leverage physician classifications to train new host response score cutoffs.

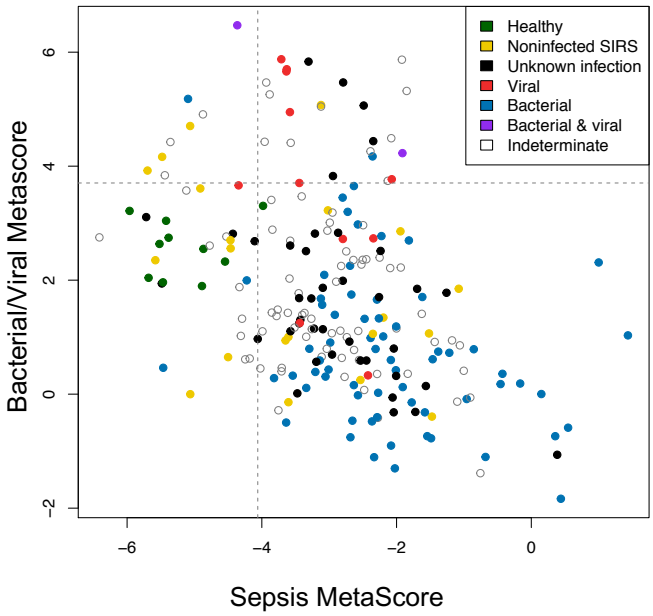
Test Cohort (n=100) Present patients' host response results with new cutoffs to physicians to assess in Main Chart Review

* 7 patients whose host response failed due to PCR amplification errors not included.
† Includes 3 patients with bacterial-fungal coinfections.

C **Example of a Patient's Host Response Result Presented to Physicians in Main Chart Review**



D



E

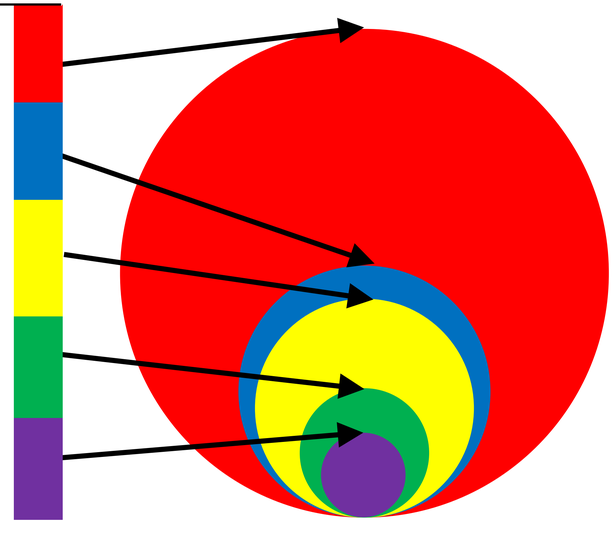
		N = 193*	Host Response: Non-Infected	Host Response: Bact-erial	Host Response: Viral
Adjudicated Physician Classification	Derivation Cohort	Noninfected	10	18	4
		Bacterial†	2	49	1
		Viral	1	4	4
	Test Cohort	Bacterial-Viral Coinfection	0	1	1
		Infected, Unknown Etiology	3	29	1
		Indeterminate	10	46	9

* 7 patients whose host response failed due to PCR amplification errors not included.
† Includes 3 patients with bacterial-fungal coinfections.

Fig. 3

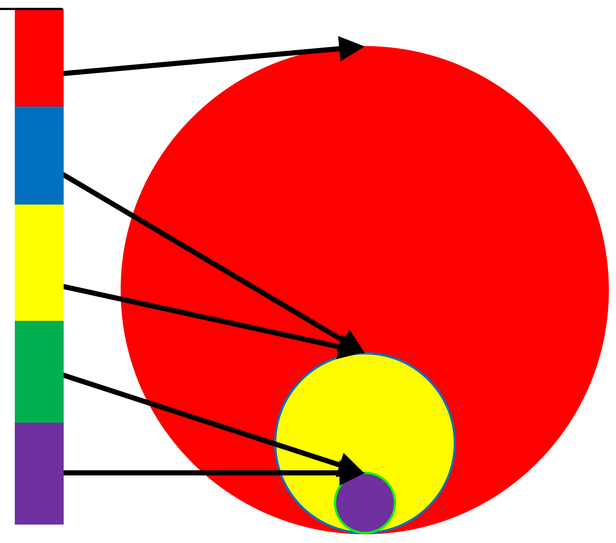
A) mNGS Clinical Utility

	Initial Physician Classification at Chart Review*					
	Noninfected	Bacterial Infection	Viral Infection	Fungal Infection or Coinfection	Probable or Unsure	All Patients
mNGS data available	16	69	11	4	100	200
mNGS results positive	4	27	2	3	17	53
Results included organism(s) not detected by standard-of-care microbiological tests within 5 days after presentation	4	17	1	2	16	40
Results included organism(s) assessed as clinically relevant by at least 2 of 3 physicians	1	8	0	0	5	14
Results led at least 2 of 3 physicians to change classification for presence of bacterial infection	1	0	0	0	5	6



B) VirCapSeq-VERT Clinical Utility

	Initial Physician Classification at Chart Review*					
	Noninfected	Bacterial Infection	Viral Infection	Fungal Infection or Coinfection	Probable or Unsure	All Patients
VirCapSeq-VERT data available	16	68	11	4	100	199
VirCapSeq-VERT results positive	2	9	4	0	12	27
Results included organism(s) not detected by standard-of-care microbiological tests within 5 days after presentation	2	9	4	0	12	27
Results included organism(s) classified as clinically relevant by at least 2 of 3 physicians	0	0	0	0	3	3
Results led at least 2 of 3 physicians to change classification for presence of viral infection	0	0	0	0	3	3

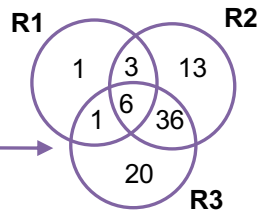


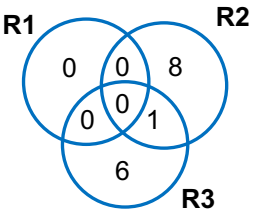
*Refers to consensus physician classifications during Phase I of Main Chart Review, when physicians were blinded to mNGS, VirCapSeq, and host response results.

Fig. 5

A

	Number of Patients for Whom Physicians Changed Classifications of Infection Status/Type After Reviewing Host Response Results		
	Reviewer 1	Reviewer 2	Reviewer 3
Classification(s) Changed ≥ 1 Point	11	58	63
Classification(s) Changed ≥ 2 Points	0	9	7





B

100 Patients Had Host Response Results Reviewed by Physicians → 25 of 100 Patients Had Positive mNGS and VirCapSeq-VERT Results → 10 of 25 Patients Had Physician(s) Change Classification for Clinical Relevance of mNGS or VirCapSeq-VERT Organisms After Reviewing Host Response Results

Patient	Host Response	Clinical Diagnosis	mNGS or VirCapSeq-VERT Organism*	Organism Clinically Relevant?					
				Reviewer 1		Reviewer 2		Reviewer 3	
				PII	PIII	PII	PIII	PII	PIII
Pt_015	Bacterial	Relapsed AML	Human herpesvirus 6 (3,550 / 20.53)	1	2	4	3	2	2
Pt_067	Viral	Pneumonia	*Propionibacterium acnes* (8, 3, 2)	1	1	1	1	2	1
Pt_071	Noninfected	Cocci Meningitis	*Enterobacter cloacae* (11, 6, 2)	1	1	2	2	3	2
Pt_095	Bacterial	Chemotherapy-Associated Fever	*Propionibacterium acnes* (27, 18, 14)	1	2	2	3	2	2
Pt_112	Bacterial	Pyelonephritis	Epstein-Barr virus (91 / 1.88)	1	1	3	2	1	1
Pt_156	Bacterial	Surgical Site Infection vs. Post-Operative Fever	Human herpesvirus 6 (21,003/30.8)	2	2	4	3	2	2
Pt_163	Bacterial	Leptospirosis	*Leptospira interrogans* (214, 186, 0)	4	4	5	5	4	5
Pt_165	Bacterial	Pneumonia vs Drug (Neulasta) Reaction	Human herpesvirus 6 (288 / 1.42)	2	2	4	3	2	2
Pt_166	Bacterial	Balanitis	*Lactobacillus mucosae* (16, 9, 5)	3	3	2	2	2	3
Pt_194	Bacterial	Radiation Pneumonitis vs Pneumonia vs Metastatic Osteosarcoma	*Haemophilus influenzae* (35, 24, 8)	3	3	4	5	4	5

* For bacterial organisms, numbers in parentheses represent raw reads, estimated lower limit for the intensity of blood-associated reads, and estimated upper limit for the intensity of contaminant reads in controls. For viruses, numbers in parentheses represent (Raw Reads / Reads per 10,000 Host Subtracted Reads).

Legend	Phase I (PI)	Phase II (PII)	Phase III (PIII)	No	Probably No	Unsure	Probably Yes	Yes
	Medical Charts Only	+mNGS, VirCapSeq-VERT	+Host Response	1	2	3	4	5