Published in final edited form as:

Sci Transl Med. 2013 September 18; 5(203): 203ra126. doi:10.1126/scitranslmed.3006280.

A Host-Based RT-PCR Gene Expression Signature to Identify Acute Respiratory Viral Infection

Aimee K. Zaas^{1,2}, Thomas Burke¹, Minhua Chen³, Micah McClain^{1,2}, Bradly Nicholson⁴, Timothy Veldman¹, Ephraim L. Tsalik^{1,2,4}, Vance Fowler², Emanuel P. Rivers⁵, Ronny Otero⁵, Stephen F. Kingsmore⁶, Deepak Voora^{1,2}, Joseph Lucas¹, Alfred O. Hero⁷, Lawrence Carin⁸, Christopher W. Woods^{1,2,4,*}, and Geoffrey S. Ginsburg^{1,2,*}

¹Institute for Genome Sciences and Policy, Duke University School of Medicine, Durham, NC 27710, USA

²Department of Medicine, Duke University School of Medicine, Durham, NC 27710, USA

³Departments of Statistics and Computer Science, University of Chicago, Chicago, IL 60637, USA

⁴Durham Veterans Affairs Medical Center, Durham, NC 27710, USA

⁵Department of Emergency Medicine, Henry Ford Health Center, Detroit, MI 48202, USA

⁶Center for Pediatric Genomic Medicine, Children's Mercy Hospital, Kansas City, MO 64108, USA

⁷Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA

⁸Department of Electrical and Computer Engineering, Duke University, Durham, NC 27710, USA

Abstract

Improved ways to diagnose acute respiratory viral infections could decrease inappropriate antibacterial use and serve as a vital triage mechanism in the event of a potential viral pandemic. Measurement of the host response to infection is an alternative to pathogen-based diagnostic

*Corresponding author. Geoffrey.ginsburg@duke.edu (G.S.G.); Christopher.woods@duke.edu (C.W.W.).

Competing interests: G.S.G., A.O.H., L.C., J.L., C.W.W., and A.K.Z. have filed for a provisional patent on the respiratory viral signature (Docket #028193-9051WO00, no. 61/181216). A.K.Z., G.S.G., C.W.W., and E.L.T. have received funding from Novartis.

Author contributions: A.K.Z. assisted in study design, interpreted the data, and wrote the manuscript. T.B. assisted in study design and analyzed the RT-PCR data. M.C. performed statistical analyses on the RT-PCR data. M.M. assisted in study design and interpretation of data. B.N. assisted in study design and performed microbiological analyses on the experimental cohorts. T.V. assisted in study design and was responsible for project management. E.L.T. performed adjudication of emergency department samples and assisted in study design. V.F., E.P.R., R.O., and S.F.K. were principal investigators on the CAPSOD study. D.V. was principal investigator on the study from which healthy controls were obtained. J.L. and A.O.H. assisted in study design and data interpretation. L.C. supervised all statistical analyses and assisted in study design, data interpretation, and manuscript preparation. C.W.W. was a principal investigator on the CAPSOD study and assisted in study design and interpretation. G.S.G. is the principal investigator and assisted in study design, data interpretation, and manuscript preparation. A.K.Z., L.C., C.W.W., and G.S.G. had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analyses.

Data and materials availability: Code for reproducing the statistical results can be found at <http://people.ee.duke.edu/~lcarin/reproduce.html>.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/5/203/203ra126/DC1

Methods

testing and may improve diagnostic accuracy. We have developed a host-based assay with a reverse transcription polymerase chain reaction (RT-PCR) TaqMan low-density array (TLDA) platform for classifying respiratory viral infection. We developed the assay using two cohorts experimentally infected with influenza A H3N2/Wisconsin or influenza A H1N1/Brisbane, and validated the assay in a sample of adults presenting to the emergency department with fever ($n = 102$) and in healthy volunteers ($n = 41$). Peripheral blood RNA samples were obtained from individuals who underwent experimental viral challenge or who presented to the emergency department and had microbiologically proven viral respiratory infection or systemic bacterial infection. The selected gene set on the RT-PCR TLDA assay classified participants with experimentally induced influenza H3N2 and H1N1 infection with 100 and 87% accuracy, respectively. We validated this host gene expression signature in a cohort of 102 individuals arriving at the emergency department. The sensitivity of the RT-PCR test was 89% [95% confidence interval (CI), 72 to 98%], and the specificity was 94% (95% CI, 86 to 99%). These results show that RT-PCR-based detection of a host gene expression signature can classify individuals with respiratory viral infection and sets the stage for prospective evaluation of this diagnostic approach in a clinical setting.

INTRODUCTION

Acute respiratory viral infections (RVIs) are common reasons for seeking medical attention. Influenza viral infection is one for which early treatment can modify disease severity and potentially prevent disease transmission (1, 2). The H1N1 influenza pandemic in 2009 highlighted the relatively poor performance of the current diagnostic armamentarium for detecting influenza, particularly in an adult population (3–8). Given that the immune response of humans differs based on the infecting pathogen class (for example, virus or bacteria) (9), we hypothesized that a host response-based diagnostic could be used to differentiate viral from nonviral etiologies of respiratory infection. Early differentiation between viral and bacterial etiology of respiratory symptoms could direct the appropriate therapy to the etiologic agent (for example, oseltamivir for influenza virus) and potentially curb misuse of antibacterial agents or improve triage in settings of a potential viral pandemic. We have previously shown that microarray profiling of host peripheral blood can accurately classify individuals with respect to several RVIs (rhinovirus, respiratory syncytial virus, influenza A H3N2/Wisconsin, and influenza A H1N1/Brisbane) (10, 11). These findings have been reproduced by others (12, 13) and validated in an independently ascertained data set (9). The objective of this study was to develop a custom reverse transcription–polymerase chain reaction (RT-PCR) assay based on the “acute respiratory viral factor” previously described (11) and to develop and validate a predictive algorithm that allows one to estimate the likelihood of a subject presenting with signs and symptoms of infection as having an RVI. Migration of the microarray-based findings to an established in vitro diagnostic platform with potential for rapid, point-of-care application represents a significant step in moving this means for diagnosing viral diseases from the realm of research to clinical use.

RESULTS

H3N2 and H1N1 influenza challenge studies in healthy volunteers are used to develop the viral infection signature

Healthy volunteers (nine men, eight women; mean age, 27 years; range, 22 to 41 years) without evidence of influenza H3N2 antibodies were included in the H3N2 cohort (Tables 1 and 2). For symptomatic participants in the H3N2 cohort [defined above (11, 14–16)], symptom onset (score ≥ 2 at a single recording) began an average of 42.7 hours after inoculation (range, 20 to 80 hours). Symptoms peaked, on average, 81 hours after inoculation (range, 44 to 104 hours; Fig. 1). Nine participants (1, 5, 6, 7, 8, 10, 12, 13, and 15) met the criteria for classification as clinically symptomatic, and of these, all nine met the criteria for classification as clinically symptomatic and microbiologically infected. For clinically symptomatic participants, the average total 5-day symptom score was 21.1 (range, 6 to 43), with an average daily peak of 6.67 (range, 2 to 11). Of the clinically asymptomatic participants ($n = 8$), two participants did not shed virus but did seroconvert at the end of the study, evidenced by a fourfold increase in influenza antibody titer at day 28. Thus, in our secondary analyses reported herein, where clinical and microbiological information was used for participant phenotyping, these two individuals (participants 2 and 3) were left out as ambiguous (fig. S1). The H1N1 challenge study was composed of 19 men and 5 women (mean age, 25 years; range, 20 to 35 years) without evidence of H1N1 antibodies at baseline. Twelve participants (2, 3, 6, 7, 8, 9, 10, 12, 13, 17, 20, and 21) met the criteria for classification as clinically symptomatic. Of these, nine participants (6, 8, 9, 10, 12, 13, 17, 20, and 21) showed evidence for microbiological infection [positive viral culture or quantitative PCR (qPCR) from nasal wash specimens taken at 48, 72, 96, or 120 hours after inoculation]. Of the clinically asymptomatic participants ($n = 12$), five participants had microbiological evidence of infection [participants 5 (seroconversion), 11, 15, 19, and 23 (positive nasal wash culture or qPCR)] (Fig. 1). Thus, in secondary analyses, where clinical and microbiological information was used for participant phenotyping, these five individuals were left out as ambiguous (fig. S2). RNA samples were not available for testing in this study from time points beyond 125 hours for analysis. For unambiguously infected participants, the average total 5-day symptom score was 17 (range, 6 to 34), with an average daily peak of 7 (range, 2 to 13). Symptoms peaked, on average, 84 hours after inoculation (range, 48 to 120 hours; Fig. 1).

An RT-PCR classifier distinguishes clinically symptomatic from clinically asymptomatic individuals with influenza

The performance of the RT-PCR assay for classification of symptomatic individuals was evaluated for H3N2 and H1N1 separately, using a training and leave-one-out cross-validation strategy on all participants in each cohort (Fig. 2). The primary phenotype analyzed was based on clinical symptom scores, represented as symptomatic (blue) or asymptomatic (red) in Figs. 2 and 3. When the classifier was trained on H3N2/Wisconsin-exposed participants, the classification error was 0/17 (0%) [area under the curve (AUC) = 1] (Fig. 2, A and C) with a cutoff score of 20. When the classifier was trained on H1N1/Brisbane-exposed participants, the classification error was 3/23 (13%) (AUC = 0.77) (2 false negatives and 1 false positive) with a cutoff score of 20 (Fig. 2, B and D). The RT-PCR

model was then trained on the cohort infected with the H3N2/Wisconsin influenza A virus and validated on the cohort infected with the H1N1/Brisbane influenza A virus (Fig. 3A). The misclassification rate was 17% (sensitivity, 90%; specificity, 76%) (Fig. 3B; AUC = 0.83). When the model was trained on the cohort infected with H1N1/Brisbane and validated on the cohort infected with H3N2/Wisconsin (Fig. 3C), the misclassification rate was 0% (sensitivity, 100%; specificity, 100%; AUC = 1) (Fig. 3D). When using participants who met the clinical and microbiological definitions of infection for cross-validation, classification error was 2/15 (13%, 2 false negatives; AUC = 0.87) for training on the cohort infected with H1N1 and testing on the cohort infected with H3N2, and 1/15 (6.7%; AUC = 0.93) when training on the cohort infected with H3N2 and testing on the cohort infected with H1N1 (figs. S3 and S4). Finally, we combined the H1N1 and H3N2 data into one total set of data and then randomly partitioned it (50% for model training and 50% for testing). We considered 100 random partitions. The mean AUC across these 100 runs was 0.975 ± 0.034 (fig. S5).

The RT-PCR viral prediction model classifies individuals with naturally acquired RVI

We tested the RT-PCR-based classifier on 28 patients with viral PCR-proven RVI [25 influenza A (22 H1N1) and 3 rhinovirus], 39 individuals with culture-proven bacterial infection (25 *Streptococcus pneumoniae*, 13 *Staphylococcus aureus*, and 1 *Escherichia coli*), and 35 healthy controls (healthy individuals enrolling in a separate study of response to aspirin) (Fig. 4 and Tables 1 and 2). To do this analysis, we used the classifier derived on the H3N2 data set above and tested on the “real-world” data set. Twenty-five of 28 (89%) patients with viral respiratory illness were classified as having viral respiratory illness, and 4 of 39 (10.3%) individuals with bacterial infection and 1 of 35 (3%) controls were classified as having viral illness (Fig. 4B). Total testing error for the group was 8% (8/102; AUC = 0.92) (Fig. 4C). Notably, two of the false negatives were from patients with H1N1 influenza infection and one was from a patient with rhinovirus infection. Of the total false positive samples, 5 of 5 (100%) had a score of ≥ 30 . Because a clinical test for RVI is unlikely to be used in healthy individuals, we also tested the data set without the use of the healthy controls. The performance of the classifier in this testing scenario was similar with misclassification of 4 of 39 (10.4%) patients with bacterial infection and only 3 of 28 patients with viral infection, resulting in a sensitivity of 89% [95% confidence interval (CI), 72 to 98%] and a specificity of 94% (95% CI, 86 to 99%) (fig. S6).

DISCUSSION

We present evidence that targeted blood-based host gene expression patterns discovered with gene expression microarrays can be successfully transitioned to an existing in vitro diagnostic platform and used prospectively for classification of infectious respiratory diseases in cohorts that are representative of typical emergency department populations. Specifically, we targeted RNA transcripts of genes identified in a microarray-derived acute RVI signature in a custom multianalyte, quantitative RT-PCR assay, using commercially available and qualified probes. The performance of the RT-PCR assay to identify symptomatic versus asymptomatic individuals with experimentally inoculated influenza H1N1 or H3N2 is excellent (sensitivity, 87 to 100%), with cross-validation highly accurate

between influenza subtypes. When tested in patients derived from several experimental cohorts, the host-based RT-PCR assay performs with a sensitivity (89%; Fig. 3) that is superior to the reported performance of rapid antigen-based diagnostics and comparable to the reported performance of viral PCR (4, 6, 8) (reported sensitivity ranges from 20 to 70%; table S5) and with comparable specificity (94%; Fig. 3) with these same commercial testing platforms (>95%; table S5). Finally, the viral infection “score” represents a measure with potential clinical benefit for both “ruling in” and “ruling out” viral respiratory infection [sensitivity, 89% (95% CI, 72 to 98%); specificity, 94% (95% CI, 86 to 98%)], with possibilities for improved triage, decreased utilization of inappropriate antibacterial therapy, and guiding judicious use of limited antiviral resources in a pandemic setting.

Using RT-PCR, we have shown that we could discriminate influenza A–infected subjects from nonsymptomatic individuals, regardless of the subtype of influenza A strain (H3N2 or H1N1) used for infection. Notably, the classifiers disproportionally represent a small number of genes from the original microarray-derived signature, with IFI27, SIGLEC1, IFI44L, RSAD2, IFI44, and ISG15 serving as the predominantly represented genes. Further studies are needed to better understand the minimal number of gene transcripts needed to provide maximal accuracy. Additional studies may also find genes given smaller weight in this model to be of importance.

The H1N1 experimental cohort was not used to derive the genes chosen for the RT-PCR assay, providing additional independent confirmation of the robustness of the assay. This testing platform maintains cross-viral specificity for the viruses represented in our sample, as shown by both the experimental challenge study comparisons and inclusion of patients with rhinovirus infection in the real-world cohort. The broad viral sensitivity of this assay may be a distinct advantage over pathogen-based antigen or PCR testing should a new influenza virus, or potentially other viral type, begin to circulate in the community. Such an advantage might have been demonstrated in the 2009 H1N1 pandemic when traditional antigen-based influenza A testing failed to detect most of the adults infected with the new H1N1 influenza virus (7), delaying appropriate treatment and identification of the causative agent. This assay may also have proved advantageous in the more recent H3N2 influenza virus outbreak in 2012 (17) or in other new influenza virus outbreaks such as the recent avian influenza A virus subtype H7N9 outbreak (18).

Validation of the RT-PCR assay derived from experimentally infected cohorts in a naturally infected population of individuals presenting with viral respiratory infections and bacterial respiratory or bloodstream infections represents an important advance in moving this technology to a standard clinical setting. We demonstrated this using a host gene expression–based predictive model derived from an experimental cohort and validated it in a real-world population presenting to emergency department doctors with symptoms of respiratory viral illness or bacterial febrile illness. With unrelated samples derived from an emergency room–based cohort and blinded testing conditions, the RT-PCR assay accurately classified >90% of individuals as having viral infection or not. Because this classifier was developed to describe host phenomena related to active RVI, it is not designed to deliver additional information about the infection status of those individuals defined as “nonviral,” although additional clinical testing could help to determine the cause of illness in those

designated nonviral. The performance of the RT-PCR classifier is similar to what we showed in our earlier work using microarrays (11), where the acute respiratory viral classifier failed to distinguish individuals with bacterial infection from those who were healthy controls.

The test results reported here are based on setting a threshold of 20 (50% probability) as a cutoff for classifying a sample as “viral” or nonviral. Although this threshold was chosen as representative of 50% probability, the performance of the classifier is quite robust. However, this set threshold may or may not be the optimal threshold (19). Because the RT-PCR assay undergoes further validation in additional cohorts, a threshold of predictive accuracy can be set based on a risk calculation (19) determined by the goals and parameters of the testing scenario (for example, minimization of false negatives and prevalence of viral infection). The optimal threshold is related to the relative cost assigned to a false positive or false negative result. In settings where missing an individual with infection is problematic (for example, pandemic situation with quarantine), one would assign a high cost to false negatives and a lower cost to false positives and set the threshold accordingly. For other scenarios, such as when the goal is to avoid giving antibacterial agents to those with viral respiratory tract infections, one would assign a high cost to false positives.

Many genes in the RT-PCR classifier have low regression coefficients as shown in tables S2 and S3, suggesting that it may be possible to significantly reduce the number of genes represented in the classifier and achieve similar results. Genes selected from the “acute respiratory factor” (11) comprise the discriminant genes in the assay. Using a smaller set of genes might allow these to be combined with a similar classifier for bacterial infection once that is under development (20). A combined viral/bacterial gene set RT-PCR host-based assay would provide the opportunity to classify a patient as having viral or “bacterial” illness, instead of viral or nonviral illness. Although the use of RT-PCR assays on purified whole-blood RNA still remains a procedure that is not suited to rapid diagnosis in a clinical laboratory setting, the current work represents an advance over classification based on gene expression microarrays. Notably, RT-PCR gene expression signatures derived from microarray analyses of peripheral blood or tumor tissue are currently in clinical use to identify patients with cardiac transplant allograft rejection (AlloMap, <http://www.allomap.com>) or to determine risk of disease recurrence in patients with breast cancer (Oncotype DX, <http://www.oncotypedx.com>), suggesting that there are pathways for clinical adoption, regulatory approval, and reimbursement for a test of this complexity.

This is not the definitive study that will be necessary for adoption of this new approach to viral infection diagnosis into clinical care, but it is an important one nonetheless. To perform the validation, we selected individuals with conclusive microbiological or serological phenotypes (versus all-comers with febrile illness) from the emergency department cohort to develop the “gold standard” for test accuracy. With the data reported here, however, it is now possible to develop a prospective validation study in a variety of clinical settings to definitively determine the test’s accuracy. We also acknowledge that we selected the best performing set of discriminatory genes from Zaas *et al.* (11) to develop the PCR predictive model. However, genomic predictors from complex RNA data sets are not necessarily unique, and there may be other models from the original gene expression data that might be

able to classify viral infection (12, 13). Last, to make a head-to-head comparison of the host-based RT-PCR model with other testing platforms (see table S5), ideally, we would compare the AUCs for the receiver operating characteristic (ROC) curves one against another. Unfortunately, the ROC data for the tests reported in the literature were not available. Nonetheless, it is worth considering that currently available commercial tests only test for one class of virus.

In summary, we have established a “proof of concept” that host expression of a relatively small set of genes, as measured by RT-PCR from blood RNA, can be used to classify viral respiratory illness in unselected individuals presenting at an emergency department for evaluation of fever. The development of this new assay and its validation in an independent real-world patient population is an important step on the translational pathway to establishing this platform for diagnostic testing. Ultimately, clinical use of this assay will likely require prospective studies to establish its clinical utility as well as the economic analyses to make the case for reimbursement.

MATERIALS AND METHODS

All human studies were approved by relevant Institutional Review Boards and conducted with informed consent and all provisions of the Declaration of Helsinki. See Table 1 for an outline of the study.

Study design

The objective of this study was to evaluate the performance of an RT-PCR TaqMan low-density array (TLDA) platform consisting of host gene probes for the classification of individuals with RVI, both in an experimental cohort and in individuals presenting to the emergency department with signs and symptoms of infection. Gene probes represented on the RT-PCR TLDA platform were derived from a previous microarray-based analysis of peripheral blood RNA from subjects experimentally inoculated with influenza A H3N2/Wisconsin or influenza A H1N1/Brisbane. Samples from patients presenting to the emergency department were selected from a larger data set and were from individuals with microbiologically proven viral respiratory infection or systemic bacterial infection. The RT-PCR assays were initially performed on the samples from the two experimental human challenge studies with influenza A for model building and refinement. The final validation was performed on samples obtained from patients in the emergency department who presented with fever and respiratory symptoms. Statistical analysis of RT-PCR data in the final validation set was performed blinded to the sample phenotype (viral infection, no infection, and bacterial infection).

Human experimental challenge cohorts

All volunteers provided informed consent and underwent extensive pre-enrollment health screening (Table 2). As previously described, we intranasally inoculated 17 healthy volunteers with influenza A H3N2 (A/Wisconsin/67/2005) (11) at Retroscreen Ltd. (London, UK) in November 2008. We also intranasally inoculated 24 healthy volunteers with influenza A H1N1/Brisbane (H1N1) at Retroscreen Ltd. in August 2009. After 24

hours in quarantine, we instilled one of four dilutions (1:10, 1:100, 1:1000, and 1:10,000) of 10 TCID₅₀ (median tissue culture infectious dose) influenza A bilaterally into the nares of participants (groups of four or five for each dilution). The virus was manufactured and processed under current good manufacturing practices by Baxter BioScience. Every 8 hours for the first 5 days after inoculation, we collected blood into RNA PAXGene collection tubes (PreAnalytiX) according to the manufacturers' specifications. We obtained nasal lavage samples from each participant daily. This sample was used for qualitative and quantitative influenza PCR to assess the success and timing of infection because the virus used would not grow sufficiently in eggs to be measured by quantitative culture. Blood and nasal lavage collection continued throughout the duration of the 6 days of quarantine. All participants received oral oseltamivir (75 mg) (Roche Pharmaceuticals) twice daily as treatment or prophylaxis at 144 hours after inoculation. All participants were negative by rapid antigen detection of a nasal wash sample (BinaxNow Rapid Influenza Antigen; Inverness Medical Innovations Inc.) at the time of discharge.

We queried patients twice daily using a modified standardized symptom score (the Jackson score, described below) for upper respiratory tract infection (15). The symptom score requires participants to rank symptoms of upper respiratory infection (stuffy nose, scratchy throat, headache, cough, malaise, and myalgias) on a scale of 0 to 3 of "no symptoms," "just noticeable," "bothersome but can still do activities," and "bothersome and cannot do daily activities." For all cohorts, symptom scores were tabulated to determine whether participants became symptomatic from the respiratory viral challenge. Participants were classified as clinically symptomatic if the sum of symptom scores totaled ≥ 6 over the first 120 hours after inoculation. A maximum total score on any day was 18. Participants were classified as clinically symptomatic and microbiologically infected if total symptom scores were ≥ 6 over the first 120 hours after inoculation and virus was detected in nasal wash samples taken at 48, 72, 96, or 120 hours after inoculation. Detection of virus in samples taken at 24 hours after inoculation was felt to represent virus from inoculation but not indicative of replicating virus, and was not included in the definition (table S1) (14–16).

Participants were classified as asymptomatic and not infected (healthy) if the symptom score was less than 6 over the 5 days of observation, and viral shedding was not documented after the first 24 hours subsequent to inoculation as above. We tabulated the standardized symptom scores at the end of each study to determine both attack rate and time of maximal symptoms. Symptom onset was defined as the time point at which a symptom score of at least 2 was achieved at two consecutive recordings. For each challenge, we collected peripheral blood at 24 hours before inoculation with virus (baseline), immediately before inoculation (before challenge), and at set intervals after challenge, to be later processed for RNA extraction and purification.

RT-PCR assay development

We used 384-well RT-PCR cards from Applied Biosystems (<http://www.appliedbiosystems.com>). Eight samples were tested per card. To select genes for representation on the card, the top-ranking genes from the acute respiratory viral factor described in (11) were compared to a list of existing exonic primer pairs, of which 29 were

available. The “first-generation” TLDA card is described in the Supplementary Methods. The “second-generation” card used in this study comprised 48 genes: 29 genes from the acute respiratory viral gene factor [described in (11)], 3 control genes, 7 genes that were shown to be temporally down-regulated in time course analysis of the H3N2 challenge gene expression data (21), and 9 additional genes randomly selected from the original H3N2 challenge data set (11) that displayed no differential expression as possible additional controls. These gene lists and probe IDs are shown in tables S2 and S3. All samples were procured into PAXGene (PreAnalytiX) tubes per the manufacturer’s instructions. RNA was extracted at Expression Analysis from whole blood with the PAXGene 96 Blood RNA Kit (PreAnalytiX) using the manufacturer’s recommended protocol. The complete methodology for the PCR experiments can be viewed in the Supplementary Methods. RT-PCR assays were run at Expression Analysis with the ABI 7900 HT System (<http://www.appliedbiosystems.com>). Raw TLDA data were loaded into RQ Manager 1.2 for determination of C_T values. Next, .sdm-Result Data files (up to 10 plates per batch) were exported from RQ Manager and loaded into RealTime StatMiner 4.1 for data normalization and analysis. Experimental Design .txt or .csv file was generated with Sample Name, Plate, and Experimental Group for each sample data set and loaded into RealTime StatMiner for C_T analysis. C_T analysis was chosen on the basis of the manufacturer’s recommendations for use of the commercially available TaqMan assays used in this study (http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_040377.pdf).

The second-generation TLDA card was used to analyze all participants from the H3N2 and H1N1 challenges at the preinoculation time point and the time of maximal symptoms after inoculation. For all studies, relative expression of each gene was quantified with the $2^{-\Delta\Delta C_T}$ method (22), with GAPDH as the sole normalizing gene (23). $2^{-\Delta\Delta C_T}$ values were log-transformed before use in the elastic net (ENet) analysis described below.

Emergency department cohorts with bacterial and viral infection (Table 1)

Bacterial infection—Patients with infection were enrolled at Duke University Medical Center (DUMC; Durham, NC) or Henry Ford Medical Center (Detroit, MI) as part of the Community Acquired Pneumonia & Sepsis Outcome Diagnostics (CAPSOD) study—a prospective, National Institutes of Health–sponsored study to develop novel diagnostic and prognostic tests for severe sepsis and community-acquired pneumonia (ClinicalTrials.gov NCT00258869) (24, 25). Patients were screened primarily during daytime weekday hours in the DUMC or Henry Ford Emergency Department between 2006 and 2007. Patients were eligible if they had a known or suspected infection and if they exhibited two or more systemic inflammatory response syndrome criteria (26). Patients were excluded if they had an imminently terminal co-morbid condition, had advanced AIDS (CD4 count <50), were being treated with an antibiotic, or were participating in an ongoing clinical trial. *S. aureus*, *S. pneumoniae*, or *E. coli* infections were diagnosed with standard blood, sputum, or urine cultures as indicated in the Duke University or Henry Ford Hospital Microbiology laboratory. Chart review of all patients was performed to adjudicate final diagnosis, including the results of microbiologic investigations such as blood culture. Patients with

suspected or confirmed polymicrobial infections were excluded from analysis. Viral testing was not routinely performed on participants in this cohort.

Viral infection—Patients with viral respiratory infection were ascertained from patients presenting to either the DUMC or Durham Veterans Affairs Medical Center (VAMC) Emergency Department as part of the CAPSOD study or to the emergency department of the Monash General Hospital in Melbourne, Australia, in 2009. Respiratory samples were tested at the Durham VAMC Research Laboratory with the ResPlex II v2.0 viral PCR multiplex assay (Qiagen; <http://www.qiagen.com>). All samples were subsequently screened for novel H1N1 with the World Health Organization (WHO)/Centers for Disease Control and Prevention (CDC) H1N1 confirmation assay. In cases of suspected novel H1N1, samples were directly tested with the WHO/CDC assay without previous ResPlex II v2.0 testing.

A description of the microbiologic classification of the viral and bacterial infection cohorts is shown in table S4. A convenience sample of peripheral blood RNA from uninfected/healthy individuals was also used for evaluation of TLDA card performance. RNA samples were obtained from healthy volunteers participating in a study of aspirin on platelet function (27). Samples were obtained before administration of aspirin. Participants reported to be in good health at the time of enrollment; however, no formal symptom score data or microbiologic analysis for viral infection was obtained.

Development of an RT-PCR–based classifier

We performed RT-PCR TLDA assays to determine gene expression in the whole-blood RNA isolated from participants at the time of maximal symptoms (or the matching time point for asymptomatic individuals) for individuals participating in the H3N2/Wisconsin and the H1N1/Brisbane experimental studies. All $2^{-\Delta\Delta C_T}$ values were log-transformed before building a classifier. For the development of all the classifiers described in this work, the data were normalized to the PCR values of the control genes (RPL30, GAPDH, and PPIA) such that data from a given individual or a single individual could be classified. An ENet classifier was developed, which is based on probit regression (24); the ENet classifier weighs each of the genes with respect to their importance for the classification task. To “regularize” the classifier, and hence improve robustness to training data/testing data mismatch, constraints are placed on the regression weights (in terms of so-called L1 and L2 regularization). Within the ENet, we simultaneously impose a weighted combination of sparseness and smoothness on the weights. This manifests a model in which only a small subset of the genes contributes significantly to the classification decision (those with large weights), and the nonzero weights have smoothness or correlation (see tables S2 and S3). In the context of this study, the model assigns weights to each probe (gene) represented on the RT-PCR card, with certain probes contributing a greater weight to the classifier based on the level of gene expression. After the classifier is built on a training data set (for example, H3N2 or H1N1), an unseen (“left out”) sample can then be tested and assigned a probability between 0 and 1 of representing an individual with viral illness. This probability is multiplied by 40 to assign a score, which is a representation of the probability of viral illness in the individual contributing that sample (for example, an individual with 100% probability of having a viral infection based on RT-PCR results would have a score of 40) (28). The

classifier was then tested on the real-world data set, classifying samples as viral or nonviral. It should be emphasized that the probability of illness that is manifested from the model is in terms of traditional probit-regression technology, widely used within statistics. However, the amount of data we have here is too limited for quantitative verification of the probability values. Nevertheless, the performance of the model on real-world data suggests its robustness.

The ENet basic model was first described by Zou and Hastie (29), and this algorithm is state of the art. However, we also note that we obtain similar quantitative results using other related statistical models, for example, the support vector machine (30) and the relevance vector machine (31). All of these algorithms have been used, and the performance reported here is very similar across all of them (that is, the quantitative performance is not sensitive to which of these modern statistical techniques are used).

To validate the performance of the RT-PCR viral signature to accurately classify individuals as being symptomatic or asymptomatic, we trained the model on one data set (either H3N2 or H1N1) and tested it on the other independently acquired data set. To test the performance of the RT-PCR classifier in a real-world setting, we used peripheral blood RNA ascertained from patients presenting to the emergency department with febrile illness. The working procedure of the ENet classifier begins with preprocessing. We first log-transform the $2^{-\Delta\Delta C_T}$ data to obtain the data matrix X of dimension $n \times p$, where n is the number of samples and p is the number of genes. We then subtract the corresponding sample (row) mean from X and append an additional column filled with value 1.0 to account for the shift (intercept) term in the regression.

ENet training is performed with the labels of the samples Z and the preprocessed data X . We train a probit regression model $Z = \text{sign}(Y)$, $Y = X \cdot \text{beta} + \text{noise}$ with an ENet prior on the weights (beta) (24). After training, we obtain the posterior distribution of the weights (classifier). Testing was then performed. For an unseen testing sample x , we do the same data preprocessing, and the predictive probability can be obtained by integrating out the classifier using the posterior distribution in the probit model. The above predictive probability is between 0 and 1, and classification decision is made with the threshold 0.5. The final “ENet score” is obtained by multiplying a constant (for example, 40) to the predictive probability. The code for reproducing the results can be found at <http://people.ee.duke.edu/~lcarin/reproduce.html>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S. Crowe and her research assistant, E. Alesic, from Monash University for acquiring samples from Monash Hospital.

Funding: Defense Advanced Research Projects Agency grant N66001-07-C-0092 to G.S.G. and National Institutes of Allergy and Infectious Diseases grant AI066569 to S.F.K. Bacterial infection samples were obtained as part of the CAPSOD study #U01 AI066569 to S.F.K. E.L.T. was supported by Award Number 1IK2CX000530 from the Clinical Science Research and Development Service of the Veterans Health Administration Office of Research and

Development. The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation or review of the final manuscript.

REFERENCES AND NOTES

1. Ruuskanen O, Lahti E, Jennings LC, Murdoch DR. Viral pneumonia. *Lancet*. 2011; 377:1264–1275. [PubMed: 21435708]
2. Silvennoinen H, Peltola V, Vainionpää R, Ruuskanen O, Heikkinen T. Incidence of influenza-related hospitalizations in different age groups of children in Finland: A 16-year study. *Pediatr Infect Dis J*. 2011; 30:e24–e28. [PubMed: 21298851]
3. Apisarnthanarak A, Mundy LM. Rapid testing for pandemic influenza A (H1N1): Diagnostic test utility and specimen source. *Infect Control Hosp Epidemiol*. 2010; 31:663–664. [PubMed: 20433350]
4. Ciblak MA, Kanturvardar M, Asar S, Bozkaya E, Yenen OS, Badur S. Sensitivity of rapid influenza antigen tests in the diagnosis of pandemic (H1N1)2009 compared with the standard rRT-PCR technique during the 2009 pandemic in Turkey. *Scand J Infect Dis*. 2010; 42:902–905. [PubMed: 20662619]
5. Cunha BA, Syed U, Stroll S, Mickail N, Laguerre M. Winthrop-University Hospital Infectious Disease Division's swine influenza (H1N1) pneumonia diagnostic weighted point score system for hospitalized adults with influenza-like illnesses (ILIs) and negative rapid influenza diagnostic tests (RIDTs). *Heart Lung*. 2009; 38:534–538. [PubMed: 19944879]
6. de la Tabla VO, Antequera P, Masiá M, Ros P, Martín C, Gazquez G, Buñuel F, Sánchez V, Robledano C, Gutiérrez F. Clinical evaluation of rapid point-of-care testing for detection of novel influenza A (H1N1) virus in a population-based study in Spain. *Clin Microbiol Infect*. 2010; 16:1358–1361. [PubMed: 21382125]
7. Keitel K, Wagner N, Lacroix L, Manzano S, Gervais A. Performance characteristics of a rapid immunochromatographic assay for detection of pandemic influenza A (H1N1) virus in children. *Eur J Pediatr*. 2011; 170:511–517. [PubMed: 20938682]
8. Lee HM, Park HK, Hwang HS, Chun MY, Pai HJ, Oh SH, Kim DA. Diagnostic value of the rapid influenza antigen test for novel influenza A (H1N1). *Scand J Infect Dis*. 2011; 43:43–46. [PubMed: 20735325]
9. Ramilo O, Allman W, Chung W, Mejias A, Ardura M, Glaser C, Wittkowski KM, Piqueras B, Banchereau J, Palucka AK, Chaussabel D. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood*. 2007; 109:2066–2077. [PubMed: 17105821]
10. Woods CW, McClain MT, Chen M, Zaas AK, Nicholson BP, Varkey J, Veldman T, Kingsmore SF, Huang Y, Lambkin-Williams R, Gilbert AG, Hero AO III, Ramsburg E, Glickman S, Lucas JE, Carin L, Ginsburg GS. A host transcriptional signature for presymptomatic detection of infection in humans exposed to influenza H1N1 or H3N2. *PLoS One*. 2013; 8:e52198. [PubMed: 23326326]
11. Zaas AK, Chen M, Varkey J, Veldman T, Hero AO III, Lucas J, Huang Y, Turner R, Gilbert A, Lambkin-Williams R, Øien NC, Nicholson B, Kingsmore S, Carin L, Woods CW, Ginsburg GS. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe*. 2009; 6:207–217. [PubMed: 19664979]
12. Statnikov A, Lytkin NI, McVoy L, Weitkamp JH, Aliferis CF. Using gene expression profiles from peripheral blood to identify asymptomatic responses to acute respiratory viral infections. *BMC Res Notes*. 2010; 3:264. [PubMed: 20961438]
13. Statnikov A, McVoy L, Lytkin N, Aliferis CF. Improving development of the molecular signature for diagnosis of acute respiratory viral infections. *Cell Host Microbe*. 2010; 7:100–101. [PubMed: 20159615]
14. Barrett B, Brown R, Volland R, Maberry R, Turner R. Relations among questionnaire and laboratory measures of rhinovirus infection. *Eur Respir J*. 2006; 28:358–363. [PubMed: 16641127]
15. Jackson GG, Dowling HF, Spiesman IG, Boand AV. Transmission of the common cold to volunteers under controlled conditions. I. The common cold as a clinical entity. *AMA Arch Intern Med*. 1958; 101:267–278. [PubMed: 13497324]

16. Turner RB. Ineffectiveness of intranasal zinc gluconate for prevention of experimental rhinovirus colds. *Clin Infect Dis*. 2001; 33:1865–1870. [PubMed: 11692298]
17. Epperson S, Jhung M, Richards S, Quinlisk P, Ball L, Moll M, Boulton R, Haddy L, Biggerstaff M, Brammer L, Trock S, Burns E, Gomez T, Wong KK, Katz J, Lindstrom S, Klimov A, Bresee JS, Jernigan DB, Cox N, Finelli L. Influenza A (H3N2)v Virus Investigation Team, Human infections with influenza A(H3N2) variant virus in the United States, 2011–2012. *Clin Infect Dis*. 2013; 57(Suppl. 1):S4. [PubMed: 23794729]
18. Yu L, Wang Z, Chen Y, Ding W, Jia H, Chan JF, To KK, Chen H, Yang Y, Liang W, Zheng S, Yao H, Yang S, Cao H, Dai X, Zhao H, Li J, Bao Q, Chen P, Hou X, Li L, Yuen KY. Clinical, virological, and histopathological manifestations of fatal human infections by avian influenza A(H7N9) virus. *Clin Infect Dis*. 2013;10.1002/0471221082
19. Van Trees, H. Detection, Estimation, and Modulation Theory, Part I. John Wiley & Sons; New York: 1968.
20. Ahn SH, Tsalik EL, Cyr DD, Zhang Y, van Velkinburgh JC, Langley RJ, Glickman SW, Cairns CB, Zaas AK, Rivers EP, Otero RM, Veldman T, Kingsmore SF, Lucas J, Woods CW, Ginsburg GS, Fowler VG Jr. Gene expression-based classifiers identify *Staphylococcus aureus* infection in mice and humans. *PLoS One*. 2013; 8:e48979. [PubMed: 23326304]
21. Huang Y, Zaas AK, Rao A, Dobigeon N, Woolf PJ, Veldman T, Øien NC, McClain MT, Varkey JB, Nicholson B, Carin L, Kingsmore S, Woods CW, Ginsburg GS, Hero AO III. Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. *PLoS Genet*. 2011; 7:e1002234. [PubMed: 21901105]
22. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-CT} method. *Methods*. 2001; 25:402–408. [PubMed: 11846609]
23. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002; 3 RESEARCH0034.
24. Glickman SW, Cairns CB, Otero RM, Woods CW, Tsalik EL, Langley RJ, van Velkinburgh JC, Park LP, Glickman LT, Fowler VG Jr, Kingsmore SF, Rivers EP. Disease progression in hemodynamically stable patients presenting to the emergency department with sepsis. *Acad Emerg Med*. 2010; 17:383–390. [PubMed: 20370777]
25. Tsalik EL, Jones D, Nicholson B, Waring L, Liesenfeld O, Park LP, Glickman SW, Caram LB, Langley RJ, van Velkinburgh JC, Cairns CB, Rivers EP, Otero RM, Kingsmore SF, Lalani T, Fowler VG, Woods CW. Multiplex PCR to diagnose bloodstream infections in patients admitted from the emergency department with sepsis. *J Clin Microbiol*. 2010; 48:26–33. [PubMed: 19846634]
26. Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, Schein RM, Sibbald WJ. ACCP/SCCM Consensus Conference Committee, Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. 1992. *Chest*. 2009; 136(Suppl. 5):e28. [PubMed: 20162763]
27. Voora D, Ortel TL, Lucas JE, Chi JT, Becker RC, Ginsburg GS. Time-dependent changes in non-COX-1-dependent platelet function with daily aspirin therapy. *J Thromb Thrombolysis*. 2012; 33:246–257. [PubMed: 22294277]
28. Chen M, Carlson D, Zaas A, Woods CW, Ginsburg GS, Hero A III, Lucas J, Carin L. Detection of viruses via statistical gene expression analysis. *IEEE Trans Biomed Eng*. 2011; 58:468–479. [PubMed: 20643599]
29. Zou H, Hastie T. Regularization and variable selection via the Elastic Net. *J R Statist Soc B*. 2005; 67:301–320.
30. Vapnik, VN. The Nature of Statistical Learning Theory. Springer-Verlag; New York: 1995.
31. Tipping ME. Sparse Bayesian learning and the relevance vector machine. *J Mach Learn Res*. 2001; 1:211–244.

Fig. 1. Symptom scores for experimental cohorts

Modified Jackson scores (15) were recorded for participants in the experimental cohorts, from preinoculation through 120 hours after inoculation. Scores were recorded twice daily, and the higher of the two scores was used to compile the aggregate score over 120 hours. **(A)** Daily modified Jackson scores for the nine participants who developed symptoms of acute respiratory viral illness in the H1N1 exposure study. **(B)** Daily modified Jackson scores for the nine participants who developed symptoms of acute respiratory viral illness in the H3N2 exposure study. Those who remained asymptomatic are not shown.

Fig. 2. Validation of RT-PCR expression in the H3N2 and H1N1 experimental cohorts
The performance of the RT-PCR assay for classification of symptomatic individuals was evaluated for H3N2 and H1N1 separately, using a training and leave-one-out cross-validation strategy on all participants in each cohort. **(A and C)** When a classifier was trained on H3N2/Wisconsin-exposed participants, the classification error was 0/17 (0%) (AUC = 1) with a cutoff score of 20. **(B and D)** When a classifier was trained on H1N1/Brisbane-exposed participants, the classification error was 3/23 (13%) (AUC = 0.77) (2 false negatives and 1 false positive) with a cutoff score of 20. When using only participants who met both the clinical and microbiological definitions of infection, classification error was 0/15 (0%; AUC = 1) and 1/15 (6.7%, one false negative; AUC = 0.93) for H3N2/Wisconsin and H1N1/Brisbane, respectively (figs. S1 and S2). The primary phenotype analyzed was based on clinical symptom scores, represented as symptomatic (blue) or asymptomatic (red). Symptomatic and microbologically infected individuals are represented with blue dots, and asymptomatic and microbologically uninfected individuals are represented with red dots. Asymptomatic but microbologically infected individuals are represented with red diamonds. Pd, probability of detection; Pf, probability of false discovery.

Fig. 3. Cross-validation of RT-PCR expression among experimental cohorts

Classification of individuals with symptomatic RVI was validated across influenza subtypes.

(A and B) When training on H3N2 participants and testing on H1N1 participants, the sensitivity of the test is 90% and specificity is 75% (1 false positive and 3 false negative results) (AUC = 0.83). (C and D) When training on H1N1 participants and testing on H3N2 participants, sensitivity and specificity are both 100% (AUC = 1). ENet score (y axis) represents the probability of having a viral infection, with a score of 20 indicative of 50% probability of detection. Symptomatic and microbiologically infected individuals are represented with blue dots, and asymptomatic and microbiologically uninfected individuals are represented with red dots. The primary phenotype analyzed was based on clinical symptom scores, represented as symptomatic (blue) or asymptomatic (red). Symptomatic but microbiologically uninfected individuals are represented with blue diamonds, and asymptomatic but microbiologically infected individuals are represented with red diamonds.

Fig. 4. External validation of the RT-PCR-based classifier in the emergency department cohort (A) An RT-PCR-based gene expression classifier trained on a cohort experimentally inoculated with the H3N2/Wisconsin influenza virus accurately classifies individuals presenting to the emergency department with viral infection (blue) and distinguishes these individuals from those presenting with Gram-positive bacterial infection (red) and healthy controls (green). (B) The sensitivity of the RT-PCR test is 89% (95% CI, 72 to 98%), and the specificity is 94% (95% CI, 86 to 99%), with a positive predictive value of 84% (95% CI, 65 to 94%) and a negative predictive value of 96% (95% CI, 89 to 99%; AUC = 0.92). (C) ENet score (y axis) represents probability of having viral infection, with a score of 20 indicative of 50% probability of detection. “Classifier weight” refers to the regression coefficient for each gene in the model. “Gene index” represents each gene probe in the 48-probe assay.

Table 1
Development and validation of the RT-PCR assay

ENet, elastic net; VAMC, Veterans Affairs Medical Center.

RT-PCR assay development	RT-PCR assay validation
Cohorts (experimental challenges)	Cohorts (emergency department)
H3N2 (UK, 2008)	Bacterial infections
H1N1 (UK, 2009)	Duke (United States, 2006–2008)
Second-generation assay [*]	Henry Ford Hospital (United States, 2006–2008)
48 genes	Viral infections
RT-PCR using second-generation assay	Duke (United States, 2009)
Development of ENet predictive model [†]	Durham VAMC (United States, 2009)
	Monash General Hospital (Australia, 2009)
	RT-PCR using second-generation assay
	Classification using ENet predictive model

^{*} Details of the composition of the second-generation RT-PCR assay are described in Materials and Methods.

[†] Details of the ENet predictive model development and its application are described in Materials and Methods as well as in the Supplementary Methods.

Table 2

Subject demographics, samples sources, and infection diagnosis.

Source	Year	Location	Number	Gender (% male)	Mean age (years)	Infection	Cohort
Experimental inoculation							
H3N2/Wisconsin	2008	UK	17	77	27	n/a [*]	Discovery
H1N1/Brisbane	2009	UK	24	47	25	n/a [†]	Discovery
Emergency department							
Viral infection							
Duke University	2009	Durham, NC	6	50	15	Novel H1N1	Validation
Durham VAMC	2009	Durham, NC	2	100	42	n = 1 novel H1N1; n = 1 influenza A	Validation
Henry Ford Health System	2009	Detroit, MI	13	31	37	Novel H1N1	Validation
Henry Ford Health System	2009	Detroit, MI	3	0	49	Rhinovirus	Validation
Monash General Hospital	2009	Melbourne, Australia	4	50	27	Novel H1N1	Validation
Bacterial infection							
Henry Ford Health System	2006–2007	Detroit, MI	30	60	57	n = 8 <i>S. aureus</i> ; n = 1 <i>Staph</i> ;	Validation
Duke University	2006–2007	Durham, NC	9	67	56	n = 21 <i>S. pneumoniae</i> n = 5 <i>S. aureus</i> ; n = 3 <i>S. pneumoniae</i> ; n = 1 <i>E. coli</i>	Validation
Healthy controls Duke University	2009	Durham, NC	35	67	30	n/a	Validation

^{*} Nine participants developed clinical symptoms.

[†] Nine participants developed clinical symptoms.