

Supplementary Information

Data Tables and Figures

- a. Supplementary Table 1- Study Populations and Illness Characteristics
- b. Supplementary Table 2- Study Populations and Illness Characteristics Based on Microbiology Categories
- c. Supplementary Table 3- Primary Analysis Group compared to Noninfected Controls
- d. Supplementary Table 4 – Pathogens identified
- e. Supplementary Table 5. Bacterial infections misclassified as nonbacterial using 4 gene risk threshold.
- f. Supplementary Figure 1- Expression profiles for the 4 genes in the signature.
- g. Supplementary Figure 2. Performance of WBC and PCT to discriminate bacterial and nonbacterial LRTI.

Supplementary Table 1- Study Populations and Illness Characteristics

Characteristic	Analysis Group N=504	Not analyzed N=507	P- value
Mean Age (SD)	61.2 (17.4)	63.1 (15.5)	0.067
Female, No (%)	281 (56)	275 (54)	0.658
Race, No (%)			0.940
White	347 (69)	344 (68)	
Black/AA	145 (29)	149 (29)	
Other	12 (2)	13 (3)	
Hispanic	53 (11)	46 (10)	0.460
Medical Conditions			
Mean BMI (SD)	31.5 (10)	30.5 (12)	0.145
Asthma	189 (38)	181 (36)	0.846
COPD	165 (33)	250 (49)	<0.0001
Any Smoking	373 (74)	423 (83)	0.0003
Home Oxygen	56 (11)	120 (24)	<0.0001
CAD	87 (17)	115 (23)	0.034
CHF	65 (13)	105 (21)	0.001
Diabetes Mellitus	152 (30)	162 (32)	0.542
Chronic Kidney Disease	56 (11)	70 (14)	0.216
Any medical condition	494 (98)	496 (98)	1.0
Mean No. Medical Conditions (SD)	3.8 (1.8)	4.4 (1.9)	<0.0001
Hospital Course			
ICU	75 (15)	53 (10)	0.038
Non-invasive Ventilation	49 (10)	53 (10)	0.754
Mechanical Ventilation	13 (3)	9 (2)	0.400
CXR – any infiltrate	166 (33)	209 (41)	0.008
In hospital death	7 (1)	4 (1)	0.384
Primary Discharge Diagnoses			
Asthma	49 (10)	51 (10)	0.916

Bronchitis	45 (9)	5 (1.0)	<0.0001
AECOPD	82 (16)	137 (27)	<0.0001
Pneumonia	100 (20)	138 (27)	0.006
Respiratory Failure	27 (5)	34 (7)	0.428
Viral Syndrome	89 (18)	41 (8.)	<0.0001
Sepsis	14 (3)	13 (3)	0.848
Other	98 (19)	88 (17)	0.417

Two sided t-test and Chi Square or Fishers exact as appropriate, unadjusted for multiple comparisons.

Supplementary Table 2- Study Populations and Illness Characteristics Based on Microbiology Categories

Characteristic	Bacterial Alone N=129	P value B vs VB	Viral/Bacterial N=95	P value VB vs V	Viral Alone N=280
Mean Age (SD)	64.7 ± 15.2	0.085	61.0 ± 16.6	0.574	59.8 ± 18.4
Female, No (%)	53 (41)	0.058	52 (55)	0.181	176 (63)
Race, No (%)		0.609		0.477	
White	98 (76)		68 (72)		181 (65)
Black/AA	30 (23)		25 (26)		90 (33)
Other	1 (1)		2 (2)		9 (3)
Hispanic	9 (7)	0.800	8 (8)	0.274	36 (13)
Medical Conditions					
Mean BMI (SD)	30.8 ± 8.8	0.495	29.9 ± 11.6	0.050	32.3 ± 9.8
Asthma	25 (19)	0.002	36 (38)	0.191	128 (46)
COPD	44 (34)	0.401	38 (40)	0.075	83 (30)
Any Smoking	99 (77)	0.511	77 (81)	0.045	196 (70)
Home Oxygen	12 (9)	0.215	14 (15)	0.356	30 (11)
CAD	18 (14)	0.576	16 (17)	0.760	53 (19)
CHF	19 (15)	0.850	13 (14)	0.593	33 (12)
Diabetes Mellitus	35 (27)	0.242	33 (35)	0.442	84 (30)
Chronic Kidney Disease	14 (11)	0.540	13 (14)	0.451	29 (10)
Any medical condition	123 (95)	0.040	95 (100)	0.576	276 (99)
Mean No. Medical Conditions (SD)	3.5 ± 1.7	0.003	4.2 ± 1.8	0.062	3.8 ± 1.8
Symptoms & Signs					
Nasal Congestion	48 (37)	<0.0001	61 (64)	0.372	195 (70)
Sore throat	29 (22)	0.011	37 (39)	0.096	138 (49)
Cough	118 (91)	0.281	91 (96)	0.746	271 (97)
Dyspnea	118 (91)	0.436	90 (95)	0.210	253 (90)
Sputum production	91 (71)	0.018	80 (84)	<0.0001	174 (62)
Feverish	89 (69)	0.885	64 (67)	0.271	170 (61)

Confusion	13 (10)	1.0	9 (9)	0.071	12 (4)
Highest pulse	108 ± 19	0.718	109 ± 22	0.082	105 ± 18.3
Lowest Systolic BP	115 ± 25	0.530	117 ± 21	0.007	125 ± 26.3
Respiratory rate	26 ± 8	1.0	26 ± 8	0.021	24 ± 7.1
Oxygen Saturation	90.6 ± 7.1	0.250	89.5 ± 7.0	0.037	91.2 ± 6.8
Temperature	38.0 ± 1.1	0.006	37.6 ± 1.0	0.364	37.5 ± 0.9
Laboratory					
WBC	15.4 ± 7.9	0.127	13.8 ± 7.5	<0.0001	8.8 ± 3.8
BUN	24.5 ± 15.8	0.176	21.5 ± 17.1	0.003	17 ± 11
PCT	7.1 ± 17.8	0.874	6.7 ± 19.6	<0.0001	0.13 ± 0.15
CXR – any infiltrate	77 (60)	0.042	43 (45)	<0.0001	47 (17)
Hospital Course					
ICU	31 (24)	1.0	22 (23)	0.0002	22 (8)
Non-invasive Ventilation	14 (11)	0.540	13 (14)	0.104	22 (8)
Mechanical Ventilation	6 (5)	0.765	6 (6)	0.001	1 (0.4)
In hospital death	4 (3)	0.398	1 (1)	1.0	2 (0.7)
Primary Discharge Diagnoses					
Asthma	0	0.0004	9 (9)	0.291	40 (14)
Bronchitis	16 (12)	0.833	10 (11)	0.267	19 (7)
AECOPD	9 (7)	0.004	19 (20)	0.882	54 (19)
Pneumonia	57 (44)	0.018	27 (28)	<0.0001	16 (6)
Respiratory Failure	9 (7)	0.245	3 (3)	0.579	15 (5)
Viral Syndrome	1 (0.8)	0.0002	12 (13)	0.003	76 (27)
Sepsis	9 (7)	0.782	5 (5)	0.001	0
Other	28 (22)	0.031	10 (11)	0.022	60 (21)

Two sided t-test and Chi Square or Fishers exact as appropriate, unadjusted for multiple comparisons.

Supplementary Table 3- Noninfected Control Participant Characteristics

Characteristic	Primary Analysis ARI N=504	Non-Infected Controls N=63	P-value
Mean Age (SD)	61.2 (17.4)	65.4 (13.7)	0.066
Female, No (%)	281 (56)	21 (30)	0.001
Race, No (%)			0.628
White	347 (69)	44 (70)	
Black/AA	145 (29)	19 (30)	
Other	12 (2)	0	
Hispanic	53 (11)	2 (3)	0.070
Medical Conditions			
Mean BMI (SD)	31.5 (9.9)	33.6 (9.4)	0.111
Asthma	189 (38)	19 (30)	0.271
COPD	165 (33)	13 (21)	0.061
Any Smoking	373 (74)	41 (65)	0.135
Home Oxygen	56 (11)	7 (11)	1.0
CAD	87 (17)	16 (25)	0.120
CHF	65 (13)	26 (41)	<0.0001
Diabetes Mellitus	152 (30)	23 (37)	0.313
Chronic Kidney Disease	56 (11)	15 (24)	0.008
Any medical condition	494 (98)	59 (94)	0.060
Hospital Course			
ICU	75 (15)	10 (16)	0.852
Non-invasive Ventilation	49 (10)	5 (8)	0.852
Mechanical Ventilation	13 (3)	1 (2)	1.0
CXR – any infiltrate	166 (33)	9 (14)	<0.0001
In hospital death	7 (1)	0	1.0
Primary Discharge Diagnoses			
Asthma	49 (10)	4 (6)	0.495
Bronchitis	45 (9)	0	0.006

AECOPD	82 (16)	3 (5)	0.014
Congestive heart failure	7 (1)	28 (44)	<0.0001
Pneumonia	100 (20)	0	<0.0001
Respiratory Failure	27 (5)	0	0.060
Viral Syndrome	89 (18)	0	<0.0001
Sepsis	14 (3)	0	0.378
Atrial fibrillation	0	7 (11)	<0.0001
Pulmonary emboli	0	18 (29)	<0.0001
Other	91(18)	3 (5)	0.006

Two sided t-test and Chi Square or Fishers exact as appropriate, unadjusted for multiple comparisons.

Supplementary Table 4. Numbers of Microbial Pathogens Identified

Pathogen	Number
Virus*	
Adenovirus	8
Coronavirus	
Seasonal undifferentiated	7
NL63	1
OC43	3
229E	2
HKU1	3
SARS CoV-2	12
Human Metapneumovirus	29
Influenza A	118
Influenza B	14
Parainfluenza	
Parainfluenza undifferentiated	11
PIV-1	2
PIV-2	1
PIV-3	18
PIV-4	1
Respiratory Syncytial Virus	49
Rhinovirus	116
Bacteria*	
Gram Positive	
Beta Streptococcus	7
Alpha Streptococcus	2
<i>Streptococcus pneumoniae</i>	59
<i>Staphylococcus aureus</i>	18
Gram Negative	
<i>Hemophilus influenzae</i>	33
<i>Hemophilus parainfluenzae</i>	2
<i>Moraxella catarrhalis</i>	14
<i>Neisseria meningitidis</i>	1
<i>Pseudomonas aeruginosa</i>	6
<i>Escherichia coli</i>	5
<i>Klebsiella pneumoniae</i>	5
Pasteurella species	1
Morganella species	1
Enterobacter species	1
Serratia species	
Other GNR	12
Atypical Bacteria	
Legionella	27
Mycoplasma	10

*Multiple viruses and multiple bacteria were detected in 4.1% and 2.0% cases, respectively

SupplementaryTable 5**Bacterial Cases Misclassified as Nonbacterial Cases**

Class	Score	Syndrome	Microbiology	Clinical Data
B	-1.5978	Pneumonia	Mycoplasma PCR +	nl WBC and low PCT, CXR atel vs pn
B	-0.8862	Pneumonia	Sputum - 1+ Klebsiella	High WBC, low PCT, CXR- atel vs pn
VB	-1.5932	Pneumonia	Flu A PCR +, Sputum-rare SP	High PCT, CXR- edema
VB	-2.1935	Pneumonia	Flu A PCR +, Urine Ag SP +	nl WBC, low PCT, CXR atel vs pn
VB	-2.1541	Pneumonia	Flu A PCR +, Urine Ag SP +	nl WBC, low PCT, CXR infiltrate
VB	-3.2315	Pneumonia	SARS-CoV-2 PCR +, Sputum - 4+ Ecoli	nl WBC, PCT, CXR- atel vs pn
B	-1.4560	Non pneumonia	Sputum - 4+ Hflu	AECOPD
B	-1.1353	Non pneumonia	Sputum - 4+ Hflu	AECOPD
B	-1.3929	Non pneumonia	Sputum - 3+ Klebsiella	AECOPD
VB	-0.9639	Non pneumonia	HRV PCR +, Sputum - 3 + Hflu	AECOPD
VB	-0.9857	Non pneumonia	HRV PCR +, Sputum - 4 + SP	Asthmatic bronchitis
B	-1.1699	Non pneumonia	Sputum - Pseudomonas	Bronchiectasis
VB	-1.7488	Non pneumonia	HRV PCR +, Sputum - 3 + SP	AECOPD
VB	-1.5740	Non pneumonia	Flu A PCR +, Sputum - 3 + SP	Asthmatic bronchitis
VB	-2.6694	Non pneumonia	Flu B PCR +, Sputum - 4+ Hflu	Asthmatic bronchitis
VB	-1.0762	Non pneumonia	RSV PCR +, Sputum - 4+ Mcat	Asthmatic bronchitis
VB	-1.5112	Non pneumonia	RSV PCR +, Sputum - 4+ SP/Hflu	AECOPD
VB	-1.9231	Non pneumonia	Flu A PCR +, Sputum - 3 + Hflu	Bronchitis
VB	-1.8936	Non pneumonia	HRV PCR +, Sputum - 4+ Hflu	AECOPD

VB	-3.5322	Non pneumonia	Flu A PCR +, Sputum - Mcat	Smoker, bronchitis
VB	-2.3505	Non pneumonia	HRV PCR +, Sputum - 4+ SP	AECOPD
VB	-1.1240	Non pneumonia	Flu A PCR +, Throat culture - GAS	Streptococcal pharyngitis
VB	-0.9609	Non pneumonia	HRV 2+ Hflu in sputum	Asthmatic bronchitis

Polymerase chain reaction = PCR, WBC =white blood cell count, PCT= procalcitonin, nl = with in normal limits

CXR = chest radiograph, atel =atelectasis, pn=pneumonia

Flu A = influenza A

Flu B = influenza B

HRV = rhinovirus

RSV = respiratory syncytial virus

SP = *S. pneumoniae*

Ecoli = *Escherichia coli*

Hflu = *Hemophilus influenzae*

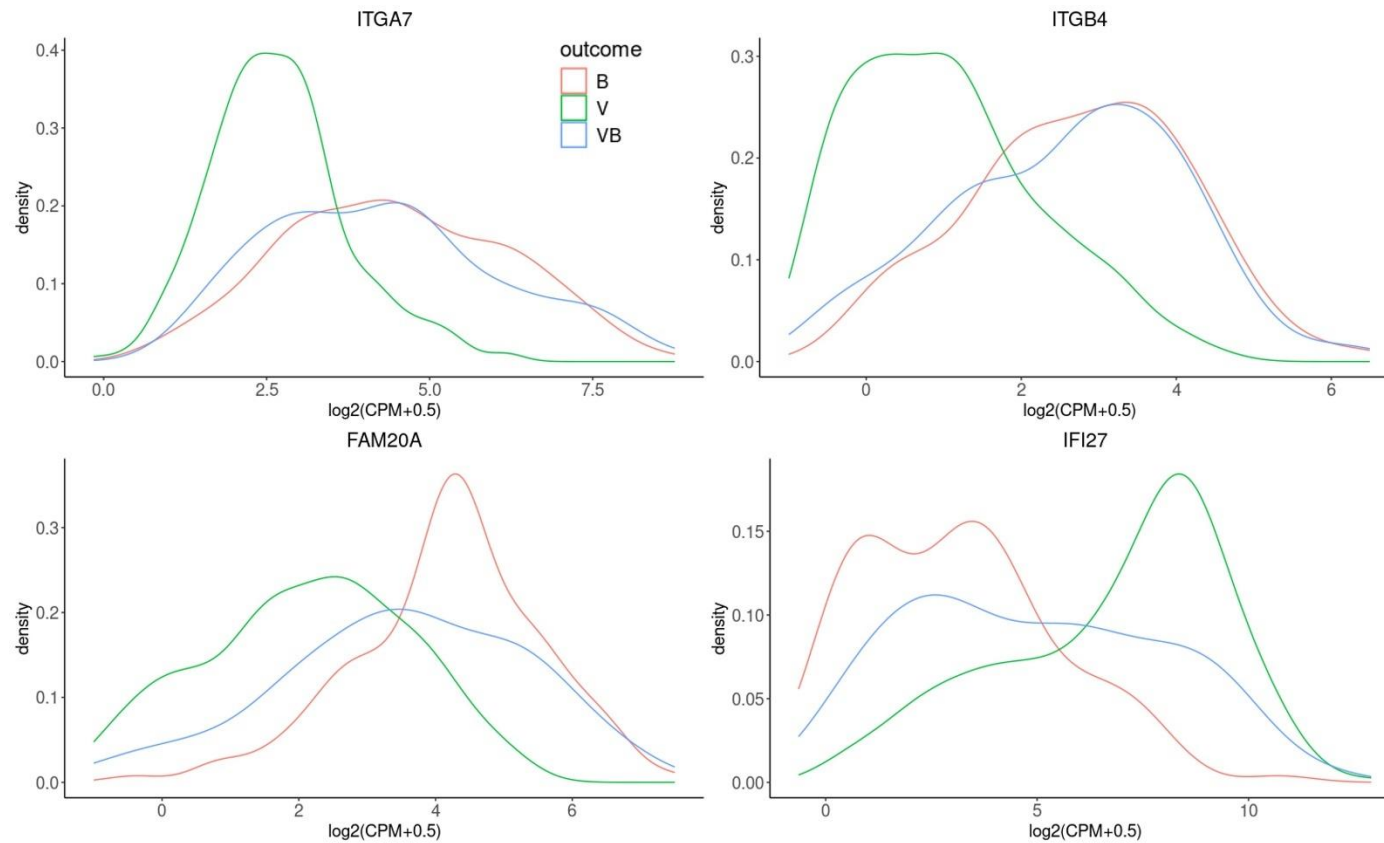
Mcat = *Moraxella catarrhalis*

GAS = Group A streptococcus

Supplementary Figure 1

Expression profiles for the 4 genes in the signature.

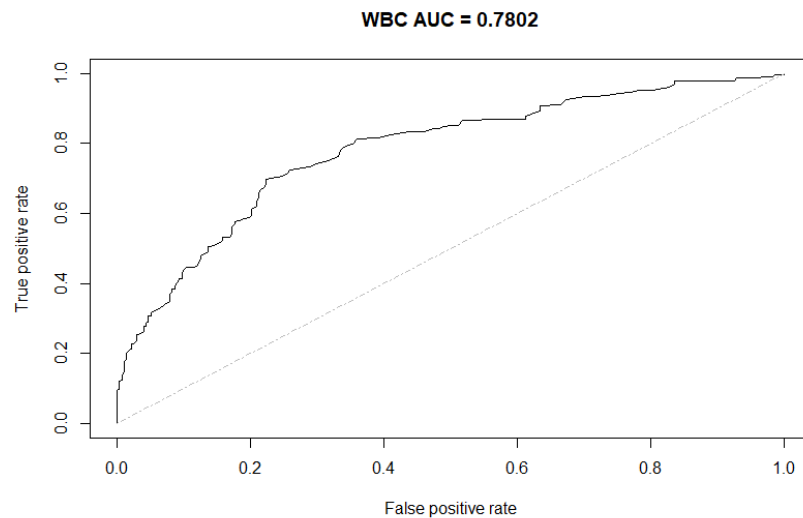
Distribution of gene expression for each of the 4 genes comprising the diagnostic gene signature by outcome. Bacterial alone (B) shown in red, Viral alone (V) shown in green and mixed Viral bacterial (VB) shown in blue. CPM=Counts per million. Source data are provided in the source data file.



Supplementary Figure 2

Performance of white blood cell count (WBC) and serum procalcitonin (PCT) in LRTI subjects to distinguish LRTI subjects with any bacterial infection (B +BV) from those with a nonbacterial (V) etiology in **Figures 2A and 2B**, respectively. The WBC (502 non-missing values) area under the curve (AUC) was 0.78 and the PCT (492 non-missing values) AUC was 0.81. Combining WBC and PCT in the same model yielded AUC = 0.86 (still lower than our 4-gene signature, despite the upward bias of this clinical model). Source data are provided in the source data file.

2A.



2B.

